DFT Study of Three-centered Hydrogen Bond in DNA Base Pairs

CHIU Lai Fan

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Master of Philosophy

in

Chemistry

December 2004


香港中文大學
大學
統系館書圖
18 NOV 2006
UNIVERSITY
LIBRARY SYSTEM
THE CHINESE UNIVERSITY OF HONG KONG

MASTER OF PHILOSOPHY (2004) The Chinese University of Hong Kong
(Chemistry)

TITLE: DFT STUDY OF THREE-CENTERED HYDROGEN BOND IN DNA BASE PAIRS

AUTHOR: CHIU Lai Fan

SUPERVISOR: Professor AU-YEUNG Chik Fun Steve

NUMBER OF PAGES: 73

Abstract

Three-centered hydrogen bonds (TCHBs) are often found in crystal and solution structures of a number of biologically relevant systems. Compared to the large amount of data available on two-centered hydrogen bonds characterization, relatively few biological systems that have been studied provide insights on the characterization of TCHBs. However, it is believed that their presence not only contributes to the flexibility of DNA, but could also serve as a possible pathway for interbase proton transfer along DNA sequence and consequently may affect its integrity. The purpose of this work is to develop a set of characterization indexes for the identification of TCHBs in DNA using a computational chemistry approach. Based on this set of characterization indexes, the effect of interbase proton transfer on the strength of hydrogen bonds is also studied.

For the computational studies, six DNA AA dimers were extracted from the solution structure of a DNA dodecamer $d(GGCAAAAAACGG)_2$. Based on the NBO analysis, a set of characterization indexes was successfully developed for the identification of TCHBs in DNA. Characterization indexes including 360 degree SAN rule, E(2) energy, Wiberg's bond index (WBI) and the scalar spin-spin coupling constant were applied. Results showed that all seven dimers conform to the 360 degree rule. The small non-zero WBI value indicates a small covalency exists in the TCHBs. Both non-zero E(2) and the scalar spin-spin coupling constant demonstrated the measure of the interaction between hydrogen atom and the distinct acceptor is present. Hence, we concluded that the presence of TCHBs can be effectively characterized by the DFT method. An exponential relationship was observed between the TCHB E(2) energy and the TCHBs distance. A positive cooperative effect was

observed for the intermolecular hydrogen bonds as the lengths of the DNA oligomer increases. Moreover, the remote control effect of interbase proton transfer on both two-centered and three-centered hydrogen bonds in trimer units has also been demonstrated.

摘要

三中心氫鍵被發現以晶體或溶液的形式存在。 相對於大量關於二中心氫鍵的研究， 對三中心氫鍵特徵鑒定的研究相對地少。 可是， 三中心氫鍵的存在不單能控制脫氧核糖核酸的柔韌性， 而且較能提供一條具潛力的小徑來促使雙螺旋之間質子的轉移， 以致最後能影響脫氧核糖核酸的完整狀態。

本論文旨在透過密度函數理論(DFT)以及自然軌道(NBO)的方法來定出一系列三中心氫鍵特徵鑒定的指標， 並根據這些指標來研究螺旋之間質子轉移對氫鍵能量的影響。

本文首先從溶液結構的雙螺旋脫氧核糖核酸 $d(GGCAAAAAACGG)_2$ 中取出 AA 雙聚物單位作出研究， 並透過自然軌道等計算方法， 成功地建立出對三中心氫鍵特徵鑒定的指標。

三中心氫鍵特徵鑒定指標包括三百六十度共和角規定， 自然軌道能量， 偉伯鍵指數和原子核-原子核的連結常數。 結果顯示， 所有 AA 雙聚物均遵守三百六十度共和角規定。 在偉伯鍵指數的參數中， 顯示出三中心氫鍵帶有少量的共價性。 偉伯鍵指數和自然軌道能量非零點的出現， 更證明氫原子和遠處的受體互相作用的現象。

最後， 本文更指出氫鍵會隨著聚成簇的增加而建立起良性合作作用， 另外雙螺旋之間質子的轉移對於二中心氫鍵和三中心氫鍵的能量發揮着遙控的作用。

ACKNOWLEDGMENTS

I would like to express my sincere thanks to my project supervisor, Professor AU-YEUNG Chik Fun Steve, for his guidance and support.

In addition, special thanks go to Chiu Wing Lok Kurtz who gave his time and insights. His suggestion greatly helped in bringing the project to completion. Thanks are also due to his technical supports in workstations and building of computer clusters. Thanks are also given to NG Chok Ki Frank for his technical support in supercomputers.

I want to thank Sze Chun Ngai Eric for his advices in the calculation of NBO and J-coupling constant.

I also would like to acknowledge the support of my tolerant father. I am very proud to have him as my father.

Finally, a special thank you is extended to my friends and labmates, who have brought joy to me in the past two years, especially Mr. Ma Nap Tak.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

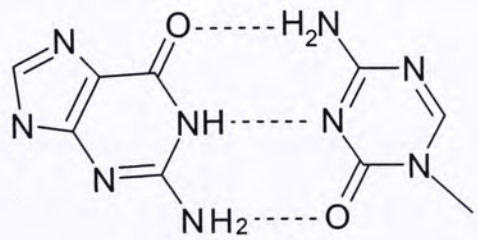
G	Deoxyguanosine
C	Deoxycytidine
A	Deoxyadenine
T	Deoxythymine
DNA	Deoxyribonucleic acid
TCHB	Three-centered hydrogen bonds
Å	Angstroms
HF	Hartree-Fock
DFT	Density Functional Theory
E	Energy
NMR	Nuclear magnetic resonance
NBO	Natural bonding orbital
NPA	Natural population analysis
E(2)	Second order perturbation energy
LP	Lone pair
BD*	Anti-bonding orbital of hydrogen bond donor
SAN	Sum of angles
WBI	Wiberg's bond index
$^1J_{XH}$	One-bond J coupling constant
Edel	Deletion energy

CHAPTER ONE

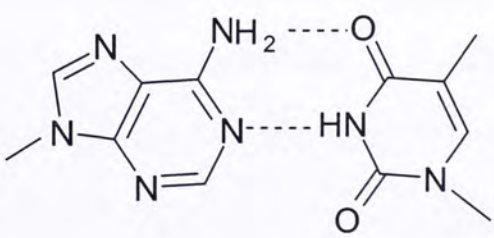
INTRODUCTION AND BACKGROUND

1.1 Introduction

Hydrogen bonding was discovered almost 100 years ago (1.1). It remains one of the most important intermolecular interactions in the biological system (1.2) and has been studied extensively (1.33-1.35). Experimental and theoretical studies of compounds with the formation of hydrogen bonds (H-bonds) showed that it can play a significant role in a wide range of chemical and biological processes (1.3). It is well known, for example, that the formation of a hydrogen bond can produce large changes in the kinetics or mechanism of genetic information propagation (1.4). On the other hand, hydrogen bond plays an essential role on holding together the two complementary strands of deoxyribonucleic acid (DNA) double helix. The two strands are linked together by hydrogen bonds formed from one donor atom to another acceptor atom between Watson-Crick base pairs. In a normal situation, G is paired up with C through three conventional hydrogen bonds whereas for the A-T base pair, there are two.



G-C base pair



A-T base pair

Currently, the nature of the hydrogen bond in DNA is of particular interest and has been explored by various experimental and theoretical methods (1.19, 1.23). In general, hydrogen bonding can be divided into several categories based on the number of proton acceptor. This chapter will briefly review the nature of two-centered H-bonds and three-centered H-bonds as well as their computational and experimental evidences.

1.2 The nature of hydrogen bonding interactions

A hydrogen bond is a kind of proton donor-acceptor interaction specifically involving hydrogen atoms. It is usually represented as A—H···B, where A is the hydrogen bond donor and B is the hydrogen bond acceptor. For the conventional type of hydrogen bond, A and B are electronegative atoms such as O and N. However, recent studies have shown, for example, that C-H group can form a weak non-conventional hydrogen bond in enzyme-substrate interactions, which is "softer" than N—H···O and O—H···O bonds (1.4).

Hydrogen bonds are suggested to form when the electronegative atom A relative to atom H in an A-H covalent bond withdraw electrons and partially deshielded one proton (1.6). At least one lone-pair or polarizable π electron(s) should be available for the proton acceptor B in order to interact with this donor A—H bond. The commonly found hydrogen bond types in biological system include O—H···O, O—H···N, N—H···O and N—H···N. An analysis of the interaction of hydrogen bond type is often decomposed into five major components including electrostatics, charge transfer, polarization, dispersion and exchange repulsion.

In general, hydrogen bonds are defined by some measurable criteria in terms of three concurring properties:

1) Shortening in the hydrogen proton-acceptor distance compared with their individual van der Waals sum distances.

2) Unusual energetic stability and,

3) the donor-hydrogen-acceptor angle should be closed to 180°.

Hydrogen bonding interaction can be sub-divided into three categories according to their energetic and geometric performance. Table 1.2.1 summarizes some common properties of strong, moderate, and weak H-bonds. It is suggested by Emsley that hydrogen bond energies which range between 15-40 kcal/mol are categorized as "strong bonds", whereas that between 1-4 kcal/mol as "weak bonds" (1.7).

Table 1.2.1 Properties of strong, moderate, and weak H-bonds following the categorization of Jeffrey (1.5)

	Strong	Moderate	Weak
Bond length H…B (Å)	1.2 – 1.5	1.5 – 2.2	> 2.2
Bond angels (°)	170 - 180	> 130	> 90
A—H versus H···B	A—H ≈ H···B	A—H < H···B	A—H << H···B
Directionality	Strong	Moderate	Weak
Bond energy (kcal/mol)[a]	15 – 40	4 - 15	< 4
Examples	HF complexes	Alcohols	Three-centered bonds

a suggested by Emsly (1.7)

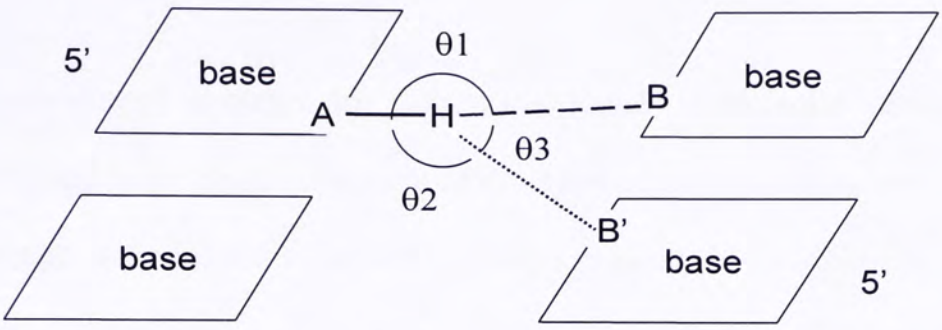
1.3 Evidences of hydrogen bonding interaction

To identify the existence of hydrogen bonding interactions in a system, three scalar quantities are commonly used. The A—H covalent bond length, the H···B hydrogen bond length and the distance between the two electronegative atoms A···B (1.6). A hydrogen bond is implicitly present when the distance between A and B is less than or equal to the sum of the van der Waals radii. For example, NMR analysis results show that the intermolecular N···O hydrogen bond in the structure of 9-methyladenine is smaller than the van der Waals radii of 2.90Å, with a measured bond distance of 2.87Å (1.14).

Experimentally, many different kinds of spectroscopic and nonspectroscopic (1.8) measurements are generally used for the determination of hydrogen bonds in biological system. Spectroscopic methods include infrared spectroscopy (IR), nuclear magnetic resonance (NMR) spectroscopy and X-ray diffraction have made rapid advances in recent years. Details of each spectroscopic method will be introduced in Section 1.4.

1.4 Three-centered hydrogen bond

Three-centered hydrogen bonds are generally considered as a long-range interaction. The group that involves hydrogen atom covalently bonded with the proton acceptor can bond to more than one acceptors B' at the same time. When the H atom simultaneously bonds with two different proton acceptor atoms B and B', a new three-centered hydrogen bond is formed (1.10). The schematic representation of a three-centered H-bond in DNA is shown in Fig. 1.4.1.



— — — two-centered H-bond

..... three-centered H-bond

Figure 1.4.1 Schematic representation of a three-centered H-bond in DNA.

For convenience, proton acceptor atoms B/B' is denoted as HB/B' in this thesis. θ1, θ2, θ3 are the three angles formed about the central H atom. θ1 is the bond angle of the two-centered H-bond which is usually close to linear for the strong type hydrogen bond, whereas θ2, θ3 can only be estimated upon the formation of three-centered H-bond. Three centered H-bonds are usually unsymmetrical with a major and minor component that r(HB') > r(HB).

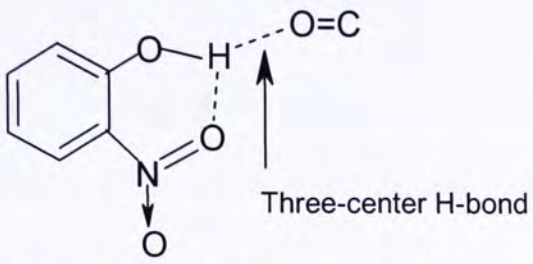
1.4.1 Literature review of three-centered H-bonds

Three-centered H-bonds are particularly found in the solid state of many compounds, and in the crystal structures of a number of biological systems (Table 1.2). Compare with the remarkable amount of data available on two-centered H-bonds, only a few model systems have been studied focusing on investigating the strength of three-centered H-bonds (1.11). The general differences between two-centered H-bonds and three-centered H-bonds are listed in Table 1.4.1.1

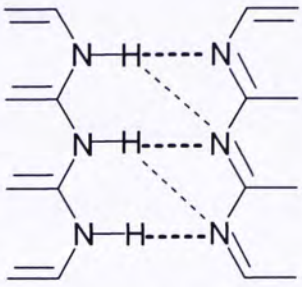
Table 1.4.1.1 Comparison between two-centered H-bonds and three-centered H-bonds

	Two-centered H-bond	Three-centered H-bond
Number of H acceptor(s)	1	2
Strength	Strong	Relatively weak
Bond length	Narrow range	Broad range
Hydrogen bond angle	Close to linear	Deviate significantly from 180°
Relative abundant in biological system (1.27)	1	> 0.25

Bureiko *el al.* have recently carried out spectroscopic studies of three-centered H-bonds of 2,6-disubstituted phenol derivatives in solution state (1.12). Results showed that the IR frequency shifts for the complexes between o-substituted phenols and CO could be elucidated by the formation of intermolecular three-centered H-bonds. Moreover, it was found that the bond strength of the original intramolecular two-centered H-bonds could be varied subsequently (1.12).



2,6-disubstituted phenol derivative



Example of DDD·AAA complex

Zimmerman and Murry have performed practical investigations of intermolecular three-centered hydrogen bonding in AAA·DDD complexes (A, D = hydrogen-bonded acceptor and donor) (1.28). Using X-ray and NMR spectroscopy methods, it was demonstrated that the formation of two-centered H-bonds occurs much more readily than the three-centered one.

Recently, ab initio calculations on some intramolecular three-centered H-bonds systems have also been undertaken. Rozas *et al* (1.29) studied the three-centered interactions in 1-phenyl-3-(2-hydroxyphenylamino)-2-buten-1-one and 1,8-diaminonaphthalene derivatives using the hybrid Density Functional Theory-Hartree-Fock method with full geometric optimization. Based on the geometry, electron density, and interaction energies analysis, they concluded that three-centered H-bonds do exist in the systems. Energetically, however, they are weaker than the normal hydrogen bonds (1.29).

Krivouruchka and coworkers (1.32) studied the complexes formed between N-(4-methyl-2-nitophenyl)acetamide and some protophilic solvents. Both the energetic and geometric parameters, including bond length, bond angle and stretching frequency have been investigated. Results show that the strength of intramolecular H-bond and the intermolecular three-centered H-bond changes in a reverse manner.

1.4.2 Review of three-centered H-bonds characterization

The existence of three-centered H-bonds has been studied extensively in both experimental and theoretical ways. Several indicators are generally used to probe their existence. From the geometrical point of view, the degree of the H atom towards the plane formed by the surrounding three heavy atoms can be investigated, which has been commonly used as a criterion for the formation of three-centered H-bonds experimentally (1.30). Broad range of three-centered H-bond length is generally regarded as a secondary way to infer the three-centered H-bonds formation due to its weak nature.

Apart from considering geometric parameters (Table 1.4.2.1), other parameters are also widely used, e.g. the interaction energy (ΔE), which is the energy difference between the cluster (E) and summation of the isolated molecules upon complexation. The energy of three-centered H-bonds is then estimated by dividing the corresponding possible number of three-centered H-bonds involved (1.13).

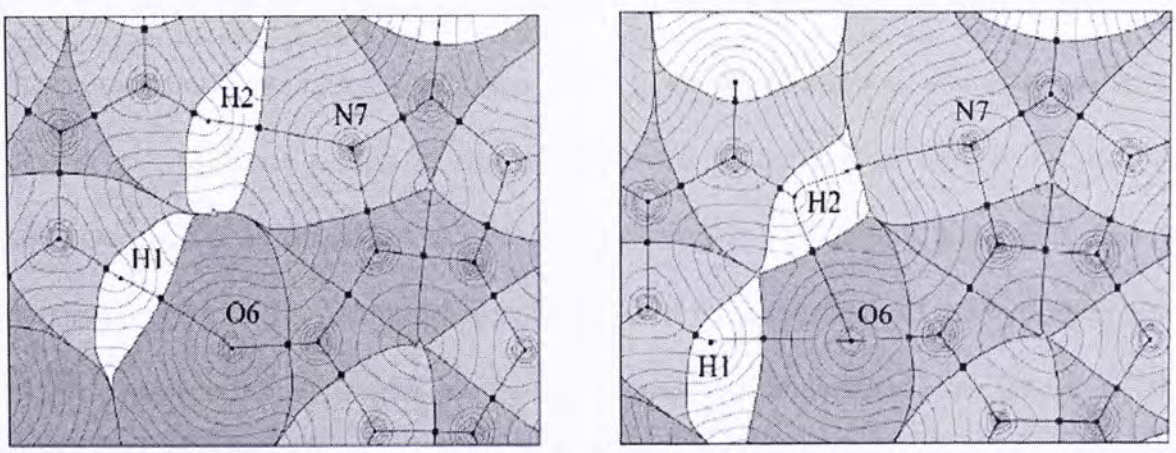
Table 1.4.2.1 Summary of three-centered H-bond geometries of some small biological molecules (crystal structure) (1.14)

Type of molecule	Type of bond	r (H···B')	θ2 (°)
Carbohydrates (neutron)	O—H···O	1.9 – 2.8	170 – 190
Purines, pyrimidines	O—H···O	1.6 - 3.0	170 – 190
	N—H···O	1.7 – 3.0	170 – 190
Nucleosides, nucleotides	O—H···O	1.8 – 2.9	170 – 190
	N—H···O	1.7 – 3.0	170 – 190

High level quantum mechanics calculations also provide an effective approach to the characterization of three-centered H-bonds in gaseous phase. With the advanced development of high speed computers, many large biological systems have been

studied theoretically. Natural Bonding Orbital (NBO) analysis accompanied with ab initio studies have broadly been chosen for quantitatively study of hydrogen bonding. Many quantum mechanic parameters can be obtained through the NBO analysis, e.g. Natural Population Analysis (NPA), Wiberg Bond Index, Milliken charge density, bond order, etc. A detail introduction will be covered in the following chapters.

Recently, topological analysis of the electron density in hydrogen bonds is also available by employing the Atoms in Molecules (AIM) theory. AIM is broadly applied to the theoretical studies of molecules in the gas phase and in small clusters (1.16). To prove the existence of hydrogen bonding of a molecule, one should find a bond path formed between the two hydrogen bond involving atoms as well as the presence of the bond critical point (BCP) in the middle of the bond path. In fact, there are eight criteria within the AIM formalism. For a detailed discussion of these criteria, the reader is referred to ref. 1.31. By measuring the distance between a BCP and a ring critical point (RCP) on the resulting molecular graph, one can estimate the stability of the hydrogen bond (1.17). However, it remains extremely difficult to apply AIM theory to the study of large cluster, such as DNA trimers.



AIM molecular graph showing the existence of BCP with (right) and without (left) the formation of three-centered H-bond (O6···H2). (1.17)

As stated earlier, it is a widespread method to identify the presence of hydrogen bond with the aid of some spectroscopic techniques, which can be applied for the characterization of three-centered H-bond as well.

For the IR spectroscopy, the proton donor A—H stretching vibration have been studied by identifying the frequencies in the absorption spectra, which are very sensitive to the formation of hydrogen bonds (1.1). Some of the IR spectral criteria for hydrogen bonding are summarized below. Experimental evidence shows that the extent of these changes actually depends on the strength of the hydrogen bonds.

(I) A—H stretching frequency moves to lower frequencies;

(II) Accompanied by an increase in intensity and band width;

(III) A—H bending frequencies move to higher frequencies;

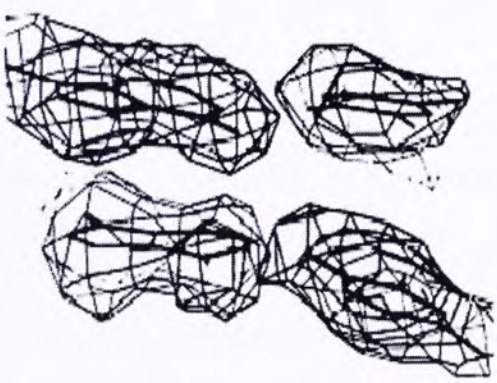
(IV) A—H stretching frequency will be varied, with increase in intensity and decrease in band width, upon cooling;

(V) Substitution of H by D lowers the A—H stretching frequencies by a factor of ~ 0.75.

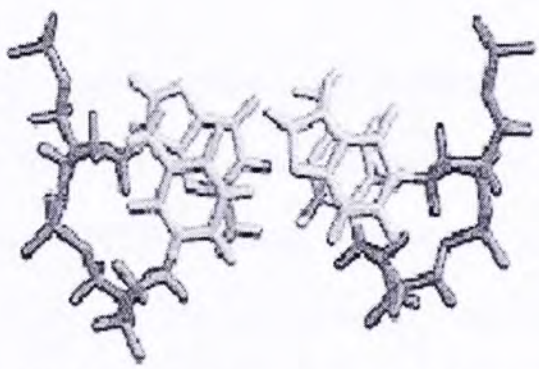
Another spectroscopic technique that has shown great potential in investigating the presence and strength of hydrogen bonds is the NMR method, which measures the degree to which the proton is shielded in terms of the proton chemical shift. The chemical shift provides proof of hydrogen bonding in both liquids and solids, and the magnitude obtained is quantitatively proportional to the strength of the hydrogen bond (1.5). Location of hydrogen atoms is essential to understand the nature of the hydrogen bond. However, hydrogen atoms cannot be observed by x-ray crystallography method, which exclusively determines the electron density

distributions of the heavy atoms. An additional simulation method is required in order to have full spatial position of the hydrogen atoms in the desire molecule. In the most accurate X-ray analyses, the H···B hydrogen bond lengths are generally too long compared with those from neutron analyses.

In contrast to X-ray crystallography technique, NMR spectroscopy can provide more precise information on the location of hydrogen atoms inside a DNA molecule. Still the results of the two techniques for hydrogen atoms often differ by more than 0.1Å (1.9). Therefore, NMR spectroscopy has been developed as a powerful tool for investigating molecular structure in the solution phase.



X-ray crystal structure of AA steps



NMR solution structure of CG base pairs

Moreover, the extension of the A—H bond upon the formation of three-centered H-bond, the intensity as well as the vibrational frequencies behavior of the H···B bond, using the spectroscopic analysis are also widely used in the literature.

1.5 Cooperative effect

One of the important phenomena associated with hydrogen bonding is the cooperative effect. In the late fifty, Frank and Wen had already suggested the presence of the cooperative effect between hydrogen bonds formed in small organic molecules (1.18). Currently, different definitions for the term "cooperativity" have been used by different authors (1.11) whose aims to explain its "contribution" on the hydrogen bonding cluster stability. In general, for example, when a hydrogen bond A—H···B forms part of a large cluster of hydrogen bonding molecules, its properties may diverge from those of the isolated A—H···B system.

A1—H1···B1 ··· A2—H2···B2 ··· A3—H3···B3 ··· A4—H4···B4

Cooperativity can be sub-divided into two categories: the positive cooperativity and the negative one. Positive cooperativity is considered as the strengthening on the first hydrogen bond when another new hydrogen bond is formed between the species involved in the previous hydrogen bond (1.11), whereas it is weakened when one has negative cooperativity. This phenomenon has already been observed in many organic and inorganic systems (1.20-1.21) using infrared spectroscopic investigations. Kleeberg (1.22) found that the cooperative effect of the OH vibrational frequency shifts in OH··· OH···B alcohol complexes can significantly increase to about 123 %. The cooperativity effect can also be revealed through other ways (1.24-1.25). For example, theoretical studies done by King and Weinhold show that HCN can exert a large cooperative effect as the (HCN)n cluster size increases, with more than 90 % energies enhancement on the corresponding hydrogen bond (1.19).

1.6 Scope of this thesis

Two-centered hydrogen bonds are believed to be the most commonly found interactions in all biological systems especially in nucleoside and nucleotides (1.14). It is believed that in nucleoside and nucleotides, three-centered H-bonds contribute more than 25 % among the hydrogen bonds found. A statistical study (1.15) of O—H···O hydrogen bond with a distance cut-off of 3.0 Å from neutron analyses of carbohydrate structure, revealed that over a quarter of normal hydrogen bonds are three-centered.

Bhattacharyya (1.16) used a fairly long distance cut-off of 3.2 Å between the hydrogen and the acceptor atoms for the determination of three-centered H-bonds in the DNA structural analysis.

Jeffery (1.5) has stated earlier that all hydrogen bonds may be characterized experimentally using classical spectroscopy techniques. These technique are however not suitable for investigating very weak three-centered H-bonds, due to the complexity and the inaccurate interpretation of the resulting spectrum. Moreover, x-ray analyses of large molecules which led to the van der Waals cut-off criterion in the A···B distance is ambiguous for the assignments of weak hydrogen bonds.

We believe that theoretical study can provide a possible alterative to address the nature of three-centered hydrogen bonding by removing the environmental perturbations. Much more attention has been focused on the role of three-centered H-bonds in the stability of biological molecules. To our knowledge, no theoretical studies have been reported on characterization of three-centered H-bonds in DNA

(nucleic acids) clusters so far. Furthermore, clear-cut correlations between three-centered H-bond energy and the three-centered H-bond geometry are currently not available. In the hope of better understanding the intrinsic character of three-centered hydrogen bonding interactions, high-level ab initio calculations with the inclusion of electron correlation effects were carried out.

In this works, Weinhold's Natural Bond Orbitals (NBO) approach was employed. The initial focus would be put on developing a set of characterization indexes for the identification of three-centered H-bonds in DNA. Later on, effort would be spent on investigating the three-centered H-bonds strength in DNA dimer, trimer and tetramer models (the clustering effect). Lastly, the position of the protons involved in the three-centered H-bonds would be varied to mimic the very beginning stage of interbase proton transfer of DNA double helix. Towards the end, we would explore the role of proton transfer on three-centered H-bonds strength.

CHAPTER TWO

THEORY AND METHODOLOGY

2.1 Introduction

Quantum mechanic (QM) calculations provide an important way of understanding the intrinsic nature of three-centered H-bonds without the presence of external perturbation such as solvent.

Practically, quantum mechanic calculation is based on the Schrodinger's equation in which electrons are represented as wave-like particles. Mathematically, these wave-like particles are described by a set of wavefunctions (2.1).

$$\text{Schrodinger's equation :} \quad H\Psi = E\Psi \tag{2.1}$$

For a multielectron system, Ψ is a many-electron wavefunction. To describe the motion of electrons, Ψ is determined by solving the Schrodinger equation. However, due to the failure in solving the many-electron Schrodinger equation, application of approximation methods are needed. To minimize the calculation variables generated from the Schrodinger equation, one of the approximations is to assume the motion of nuclei is constant and the many-electron wavefunction can be replaced by a product of one-electron wavefunctions. This approximation is then termed a Hartree-Fock (HF) approximation.

HF neglects the electron motions correlation and treats each electron separately (2.2). However, inclusion of electron correlation is vital to obtaining accurate values of nucleobase energy. Therefore, more complicated electron

correlation "post-HF" methods are employed in order to achieve more accurate theoretical results.

In recent years, a new class of QM method called Density Functional Theory (DFT) has been successfully developed to cope with the limitation of the HF method. Many earlier studies (2.3-2.4) have demonstrated that DFT methods are suitable and necessary for the structure optimization of hydrogen bonding molecules, especially in DNA. Moreover, structural parameters computed from DFT method give close agreement with the experimental structural data.

To study three-centered hydrogen bonding in DNA molecules, Natural Bonding Orbital theory (NBO) is suggested to be applied. NBO describes the formation of A−H...B hydrogen bond as the charge transfer from the lone pair of the proton acceptor to the vacant A-H antibonding orbital. The energy obtained from NBO analysis can be used as a quantitative way to estimate the relative strength of hydrogen bond.

A brief introduction of each computation method used in this study will be presented in this chapter. A more detailed discussion of methodology will also be given where appropriate.

2.2 Theory

2.2.1 DFT

The HF energy is the single-determinant eigenvalue of the electronic Hamiltonian. The electron-electron interactions integrals resulted form Coulombic

and exchange are neglected during calculation. Instead of focusing on wavefunctions and orbitals, DFT focuses on the electron density. It includes an approximate treatment of electron correlation and therefore is expected to be more accurate than Hartree-Fock theory.

In Hartree-Fock theory the energy of a system is given as follows (2.12):

$$E_{HF} = V + \langle hP \rangle + \frac{1}{2}\langle PJ(P) \rangle - \frac{1}{2}\langle PK(P) \rangle \quad (2.2)$$

where

V is the nuclear-nuclear repulsion energy,

$\langle hP \rangle$ is the one-electron kinetic and potential energy,

$\frac{1}{2}\langle PJ(P) \rangle$ is the classical electron-electron coulomb repulsion energy, and

$\frac{1}{2}\langle PK(P) \rangle$ is the exchange energy resulting from the quantum nature of electrons.

In principal, the energy based on density functional theory includes the same terms as the HF energy, which includes the nuclear, core and Coulomb terms. However, the last term in Hartree-Fock theory is replaced by two functionals, the $E^X[P]$ and $E^C[P]$ functionals. $E^X[P]$ is the exchange functional, which describes the non-classical electron-electron exchange energies of the electrons while $E^C[P]$ is the correlation functional that describes the correlated movement of electrons of different spin.

Different approaches exist for the calculation of the exchange and correlation energy terms in DFT. Apart from using the pure DFT methods, hybrid functionals are

also available. Hybrid functionals consider a mixtures of DFT exchange-correlation and Hartree-Fock exchange energies as means to improve performance.

In this study, a specific density functional theory – Becke's three-parameter hybrid functional (B3) with the 1991 nonlocal correlation functional of Perdew and Wang (PW91) is used (2.5). Becke-3-PW91 uses a different mixing scheme involving three mixing parameters provided by Becke:

$$E_{XC} = (1\text{-}0.8)*Ex(HF) + \mathbf{0.8}*E_X(LSDA) + \mathbf{0.72}*DE_X(B88) + \mathbf{0.81}*E_C(PW91) + 0.19*E_C(VWN) \quad (2.3)$$

where Local Spin Density Approximation (LSDA) is the local exchange Slater functional and the non-local correlation is provided by the Perdew 91 expression. The three highlighted constants are those determined by Becke and as used in B3LYP functional.

Our selection of B3PW91 is based on the same bonding nature between our system and Barfield's (2.4), which has been demonstrated to be satisfactorily accurate in the study of DNA structures. Moreover, previous study done by our group has shown that B3PW91 is appropriate for the calculation of hydrogen bonding interaction in DNA systems. Nevertheless, B3PW91 has its merit in computational costs and accuracy for the calculation of such systems.

To describe the hydrogen bonding system accurately in our work, 6-311G(d,p) polarization basis set has been used for the partial optimization.

2.2.2 Natural Bonding Orbital (NBO)

Many methods have been used to analyze the energy of hydrogen bonding system, such as determining the interaction energy before and after complexation. Recently, much more efforts have been focused on evaluating the delocalization contribution of hydrogen bonding interaction, which is based on the Natural Bond Orbital (NBO) approach.

NBO program performs the analysis by transforming a given wavefunction into localized form corresponding to the one-centered and two-centered elements which represent the "lone pair" and "bond' elements respectively. Generally, the wavefunction is obtained based on the optimized structure. The NBOs are obtained as local block eigenfunctions of the one-electron density matrix having optimal convergence properties for describing the electron density. The set of high-occupancy NBOs corresponds to the "natural Lewis structure" of the molecule.

NBO analysis comprises a suite of methods, including the determination of natural atomic orbitals (NAOs), natural hybrid orbitals (NHOs), natural bond orbitals (NBOs), and natural localized molecular orbirals (NLMOs). The analysis is carried out by making use of the input basis set. The occupancies of this localized basis set, which is complete and orthonormal, are mainly condensed in few areas. The Lewis-like bonds and lone pairs are likely to be found in the NBO case (2.7).

NBO analysis provides a way to transform the canonical molecular orbitals in to orthonormal set of one and two-centered localized orbitals, which are analogous to traditional Lewis-type orbitals. Currently, there are two methods available in NBO

analysis for the estimation of the relative strength of intermolecular and three-centered hydrogen bonding interaction. One method is to estimate the energy lowering effect caused by the charge transfer interaction, and the other method is to determine the deletion energy, E_{del}. E_{del} can be found by deleting the specific off-diagonal <aFb> matrix elements of the effective one-electron Hamiltonian in the NBO input basis, followed by the recalculation of the SCF energy for a hypothetical system without such orbital interactions. And the loss of the stabilization energy is E_{del}. In this study, focus will be centered on investigating the energy lowing effect caused by the charge transfer interaction.

NBO also makes use of the availability of 1-electron effective energy operator (Fock matrix) for the system. Estimation of bond energy can be obtained based on second-order perturbation theory. Two different classes of orbitals are divided during the NBO algorithm transformation, the high-occupancy "Lewis-type" orbitals and the low occupancy "non-Lewis-type" orbitals. Core orbitals, including unhydridized core-type NAO and valence lone pairs (LP) are referred to the "Lewis-type" orbitals; while σ^* and π^* antibond belong to the latter orbital type (2.8). In general, NBO theory describes the formation of a A—H···B hydrogen bond as the charge transfer from the lone pair of the base B, nB, into the higher energy level, vacant antibonding orbital σ^*(AH). An energy parameter, E(2), can be obtained by looking at the remote nB $\rightarrow \sigma^*$(AH) delocalization effect from the second order perturbation analysis. (Figure 2.2.2.1)

$$\Delta E^{(2)}_{i\rightarrow j} = \frac{-2\left\langle \delta_i \left| \hat{F} \right| \delta_j^* \right\rangle^2}{\varepsilon_j^* - \varepsilon_i} \qquad (2.5)$$

where $\hat{F}$ is the effective orbital Hamiltonian $\varepsilon_i = \langle \delta_i | \hat{F} | \delta_i \rangle$ and $\varepsilon_j^* = \langle \delta_j^* | \hat{F} | \delta_j^* \rangle$ are the respective orbital energies of donor and acceptor NBOs.

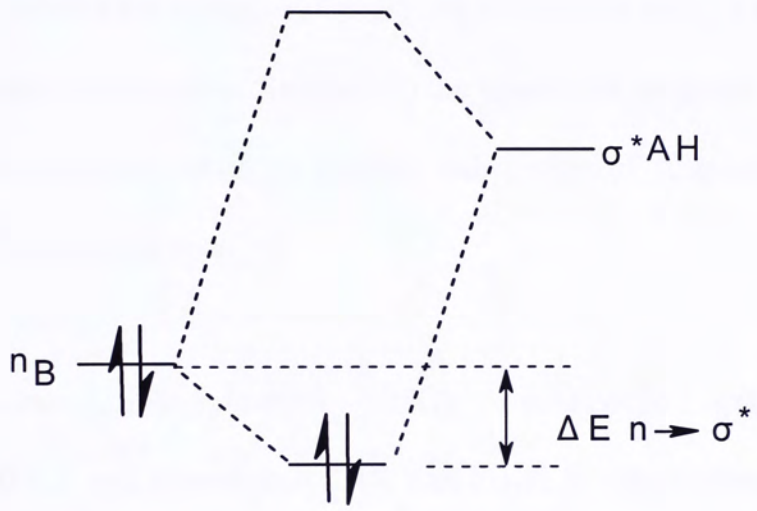


Figure 2.2.2.1 Schematic diagram shows the energy lowing for the interaction between a filled bond n_B and unfilled anibond σ^*AH

2.2.3 Spin-Spin Coupling Constants

The theory of electron-coupled interactions between nuclear spins in a molecule proposed by Ramsey was used to obtain more information about the nuclear-spin coupling constants. According to Ramsey (2.9), three different types of interaction are involved via the electron cloud of the molecules: 1) the magnetic interaction between the magnetic dipoles of the spinning electron and the nuclear spin, 2) an orbital-dipole interaction between the magnetic field and 3) the interaction between the electron and nuclear spins, the Fermi-contact. In general, calculation of coupling constants J is based on the Fermi-contact term.

The paramagnetic spin-orbit (PSO), diamagnetic spin-orbit (DSO), Fermi-contact (FC), and spin-dipole (SD) are found to contribute to the coupling constant J. Both the FC and the SD term are related to a spin polarization, while the other two terms are associated with orbital currents. However, it has been proved by Del Bene (2.10) that the Fermi-contact term strongly dominates the J-coupling. The overall contribution of the Fermi-contact term on the coupling constant is an order of magnitude larger than any other aforementioned term (2.10).

In this study, finite perturbation theory (FPT) accompanied with density functional theory (DFT) was used.

K_{XH} is the reduced isotropic coupling constant between atoms X and H and the expression for the coupling constant K_{XH} is

$$K_{XH} = (\frac{8}{3}\beta)^2 \pi \lambda^{-1} \cdot \sum_{\mu\nu} \rho_{\mu\nu}(\mu_H) < \phi_\mu | \delta(r_H) | \phi_\nu > \qquad (2.6)$$

where

β is the Bohn magneton,

λ is the perturbation parameter which indicates the perturbation added,

μ_H is the spin-density matrix,

$\delta(r_H)$ is the Dirac-delta function representing the interaction between electron and nucleus, μ and ν are the atomic orbitals.

The value of K_{XH} in equation 2.6 can be obtained by using the program implemented in Gaussian. The input keyword for calculation is represented as F(M)N form where M designates a multipole, and F(M) designates a Fermi contact perturbation for atom M.

The relation of the isotropic coupling constant K_{XH} to the usual value J_{XH} is

$$J_{XH} = (\eta / 4\pi^2) \gamma_H \gamma_X K_{XH} \qquad (2.7)$$

where γ_X and γ_H are the nuclear magnetogyric ratios for nuclei X and H, respectively.

2.2.4 Wiberg bond index

The Wiberg bond index (WBI) is a measure of the bond order based on the natural bond orbital analysis. (2.7). Unlike the concept of bond order and indices which were commonly used in classical theory of valence, the quantity of WBI is generally regarded as an "overlap" population. Through the computation of WBI values, the covalency of the chemical bond can be approximately revealed. Although they do not have values close to one and two for the single and double chemical bonds, it has been shown analytically that the WBI coincides with the 'chemist's bond order' for diatomic systems (2.11). In contrast to the Coulson MO bond order, Wiberg's index is intrinsically positive with no distinction between the net bonding or anti-bonding character of the density matrix elements.

The WBI of a hydrogen bond X...H can be represented as follow:

$$WBI_{XH} = \sum \sum \rho_{pq}^{2} \qquad (2.8)$$

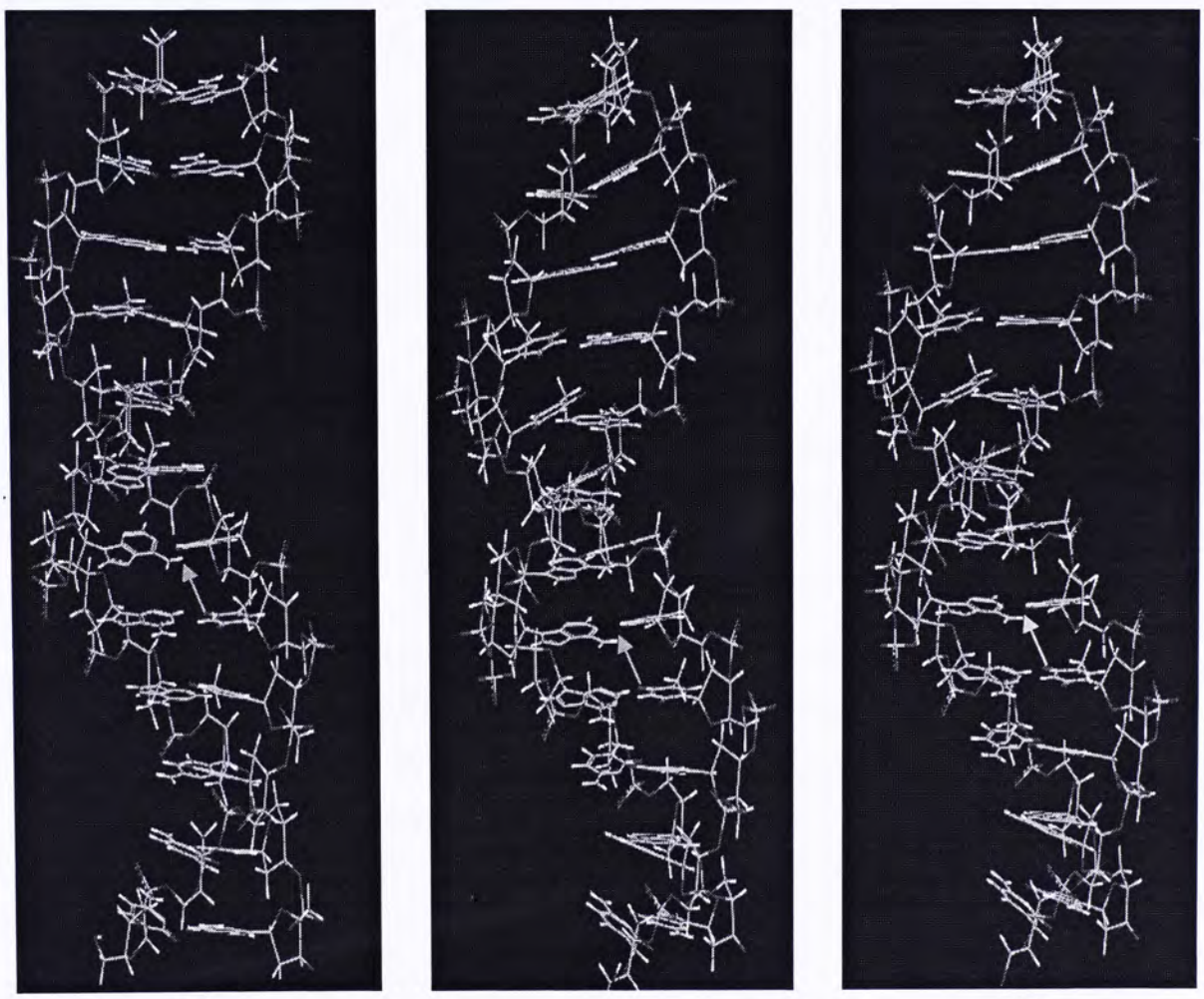
2.3 Methodology

Two DNA dodecamer sequences d(5'-G1 G2 C3 A4 A5 G6 A7 A8 A9 C10 G11 G12-3')$_2$ and d(5'-G1 G2 C3 A4 A5 A6 A7 A8 A9 C10 G11 G12-3')$_2$, which were studied in detail by solution NMR spectroscopy (Figure 2.3a, 2.3b) were obtained from Protein Data Bank (PDB codes: 1G14 and 1FZX respectively). To focus the study on hydrogen bonding interaction of the interior part of DNA duplex, all the water molecules in the NMR structure were removed.

(a) (b)

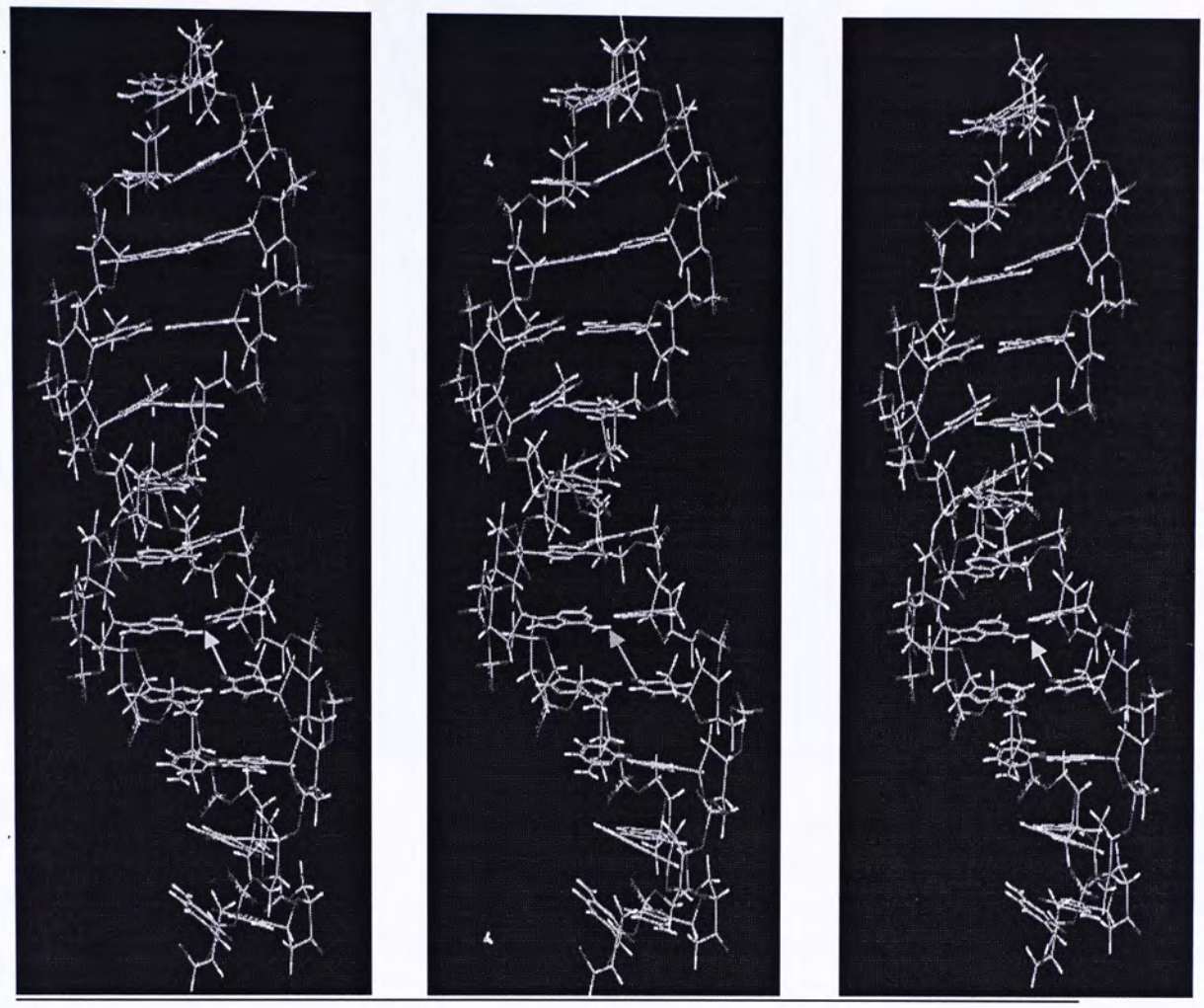
Figure 2.3a, 2.3b Superpositions of the ten best structures of 1FZX (left) and 1G14 (right)

Seven DNA A8A9 dimers from the above family of ten structures were constructed and their coordinates were extracted from the corresponding DNA dodecamers. For the purpose of this study, A8A9 dimer truncated from sequence A is denoted as "A Dimer" and that truncated from sequence B is denoted as "B Dimer" and *vice versa* (Figure 2.3.1a-g). The full A8A9 dimer unit with the individual hydrogen bonds were labeled and depicted in Figure 2.3.2.



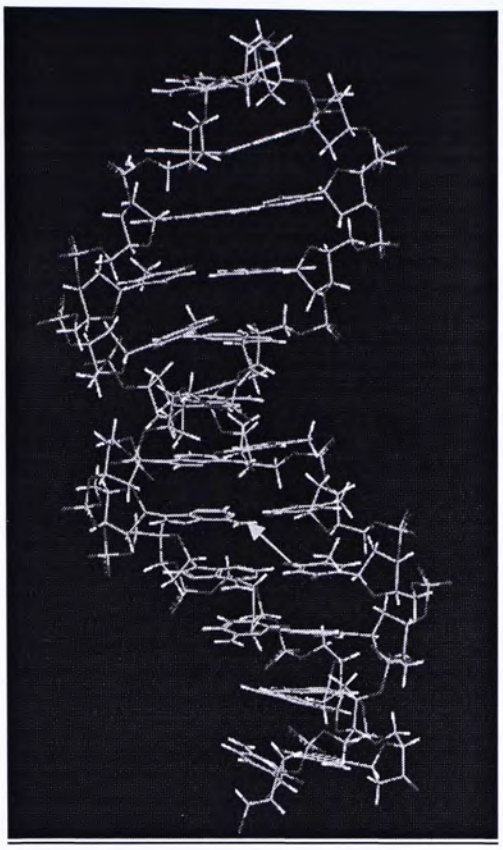
(A) (B) C)

Figure 2.3.1 a-c The NMR structures of the family of ten sequences. Sequence A (left), sequence B(middle) and sequence C (right). The arrows indicating the location of three-centered H-bond of A8A9 dimer



(D) (E) (F)

Figure 2.3.1 d-f The NMR structures of the family of ten sequences. Sequence D (left), sequence E (middle) and sequence F (right). The arrows indicating the location of three-centered H-bond of A8A9 dimer



(G)

Figure 2.3.1 g The NMR structure of the family of ten sequences - sequence G
The arrows indicating the location of three-centered H-bond of A8A9 dimer

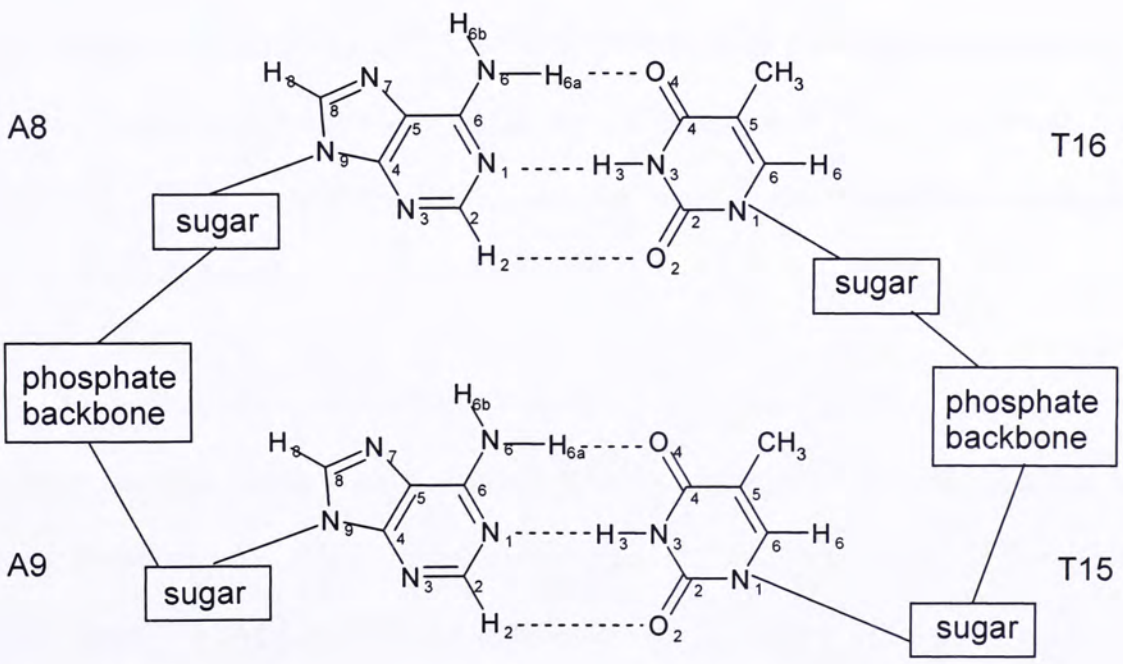


Figure 2.3.2 Schematic representation of A8A9 dimer

All *ab initio* density functional results were calculated using the GAUSSIAN 98 suite of computer programs on our parallelized cluster Pentium 3 and Sun64 Orpin computers using LINDA. In order to neutralize the negative charged state of the phosphate backbone, two protons for each dimer unit had been added. Partial optimization for the added protons dimers were then carried out at the hybrid DFT calculations using the UB3PW91/6-311G** level of theory. The so optimized geometries were then used to compute the single point energy at the same level of theory.

Natural bonding orbital (NBO) analysis - the lowering effect of the second order perturbation energy, E(2) energy, Fermi contact term of J-coupling constant, $^{1h}J_{HX}$ and 360° SAN rule were performed using the same level of theory. All quantum chemical calculations have been performed using Gaussian 98.

To investigate the effect of interbase proton transfer on the three-centered H-bond strength. Coordinates of hydrogen atoms involving in three-centered H-bonds of A7A8A9 trimer were adjusted. Distances between A-H were adjusted from 0.91 Å to 1.31Å (both bond lengths are 1.01Å at equilibrium) using 0.1Å increment each. NBO calculations were then carried out for each of the bond-distance-adjusted A7A8A9 trimer models.

To explore the cooperative character of hydrogen bonding in large DNA oligomer, models which consisted of three to five base pairs were studied. For the trimer model (2.3.3), A7A8A9, tetramer model (2.3.4), A6A7A8A9, and pentamer (2.3.5) model, A5A6A7A8A9, the coordinates of the added hydrogen atoms were partially optimized at the UB3PW91/6-311G** level. NBO analyses were then carried out followed by the single point energy calculation.

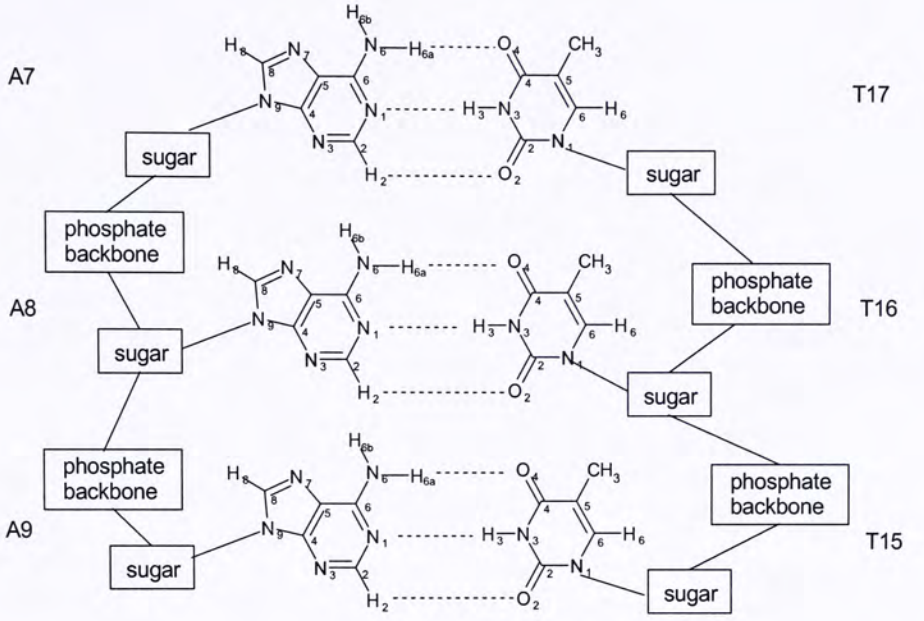


Figure 2.3.3 Schematic representation of A7A8A9 trimer

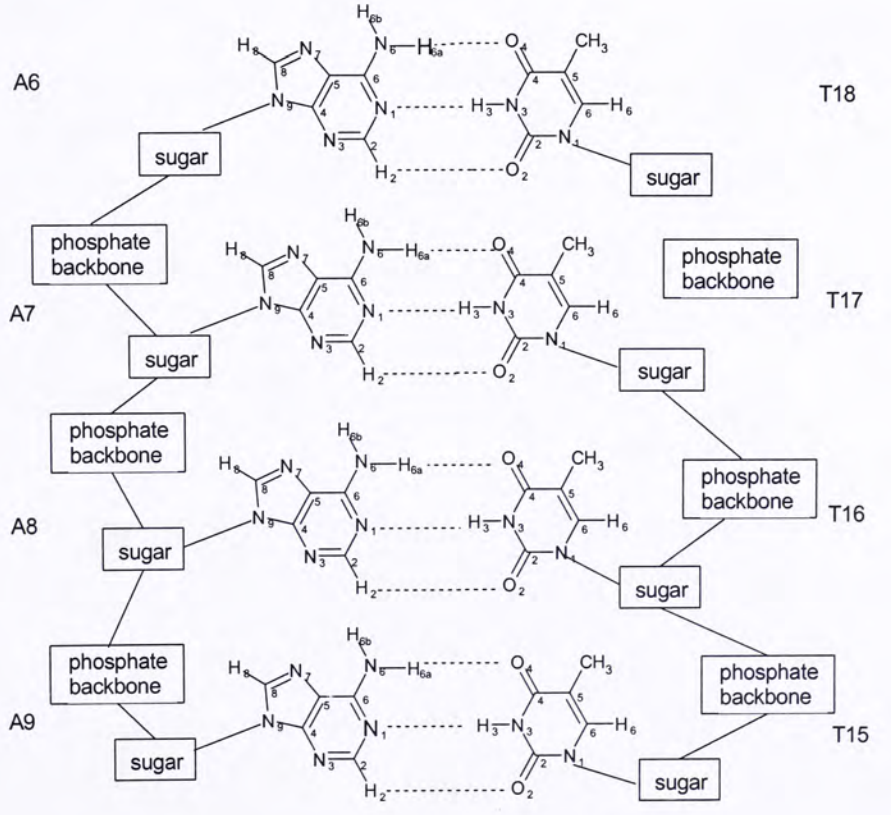


Figure 2.3.4 Schematic representation of A6A7A8A9 tetramer

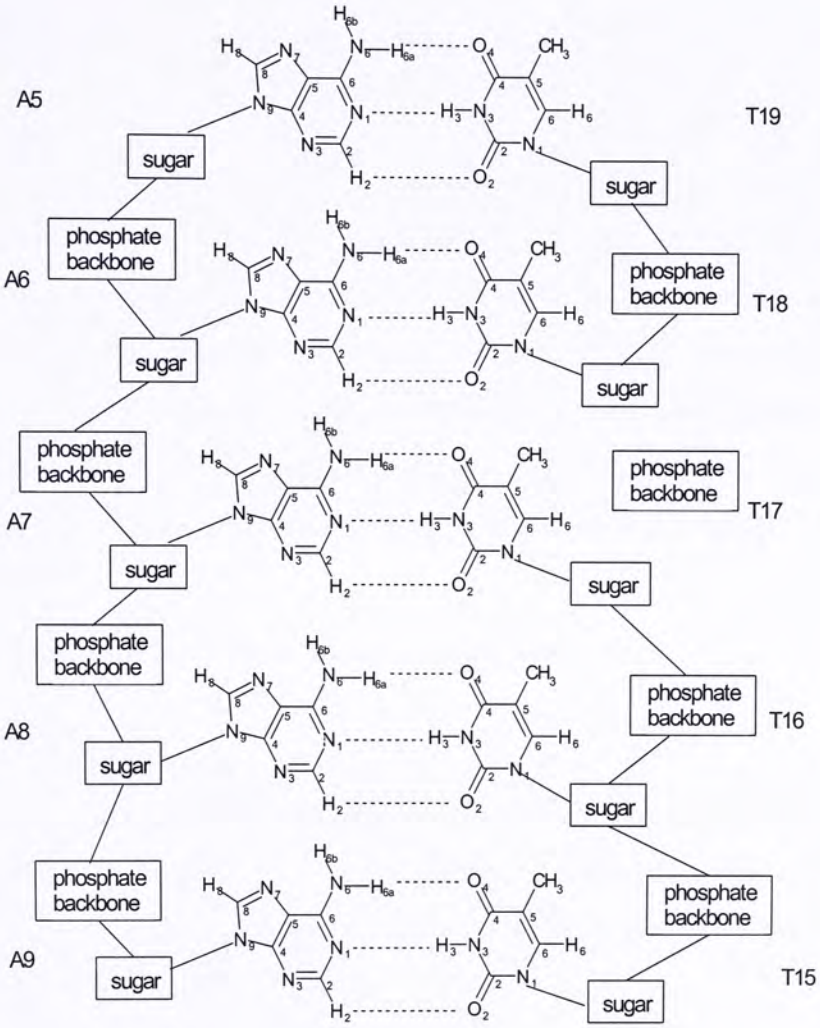


Figure 2.3.5 Schematic representation of A5A6A7A8A9 pentame

CHAPTER THREE
RESULTS AND DISCUSSION

In the previous chapter, detailed introduction of computational theory and quantum chemistry calculation methods have been discussed. In the following sections, a series of characterization indexes such as SAN rule, second-order perturbation E(2) energy, spin-spin coupling constants as well as Wiberg bond index (WBI) for the identification of three-centered H-bonds will be examined. Afterward, energetic consequences of transferring a proton toward its proton acceptor and the cooperative effect found in different size of $(AA)_n$ oligomers will be discussed.

3.1 The Nature of three-centered hydrogen bond interactions

3.1.1 Geometries

The systems studied with the formation of three-centered H-bonds are represented in Figures 3.1.1.1-3.1.1.7. The bond distances and angles of the intermolecular, two-centered hydrogen bonds and the three-centered H-bonds are summaried in Table 3.1.1.1.

As partial optimizations are carried out for the hydrogen atoms that are added on the phosphate backbone, most of the geometrical parameters of base pairs such as bond lengths and bond angles are not notably altered. In short, they are basically identical to values obtained from the experimental NMR structure.

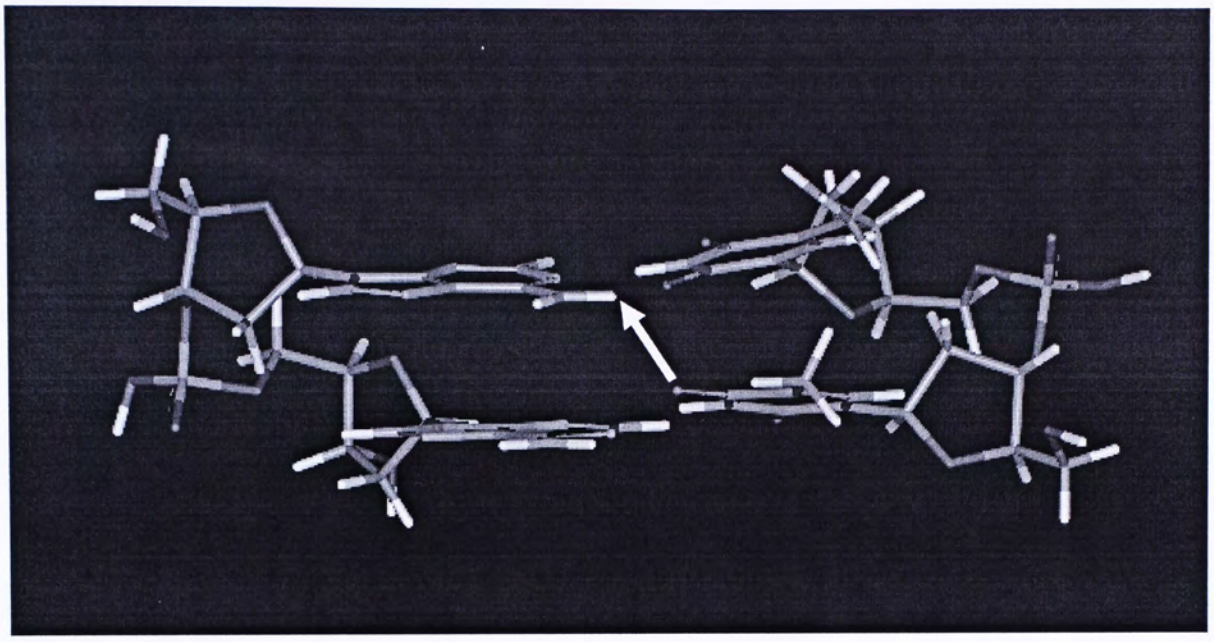


Figure 3.1.1.1 Three-centered H-bond in A Dimer

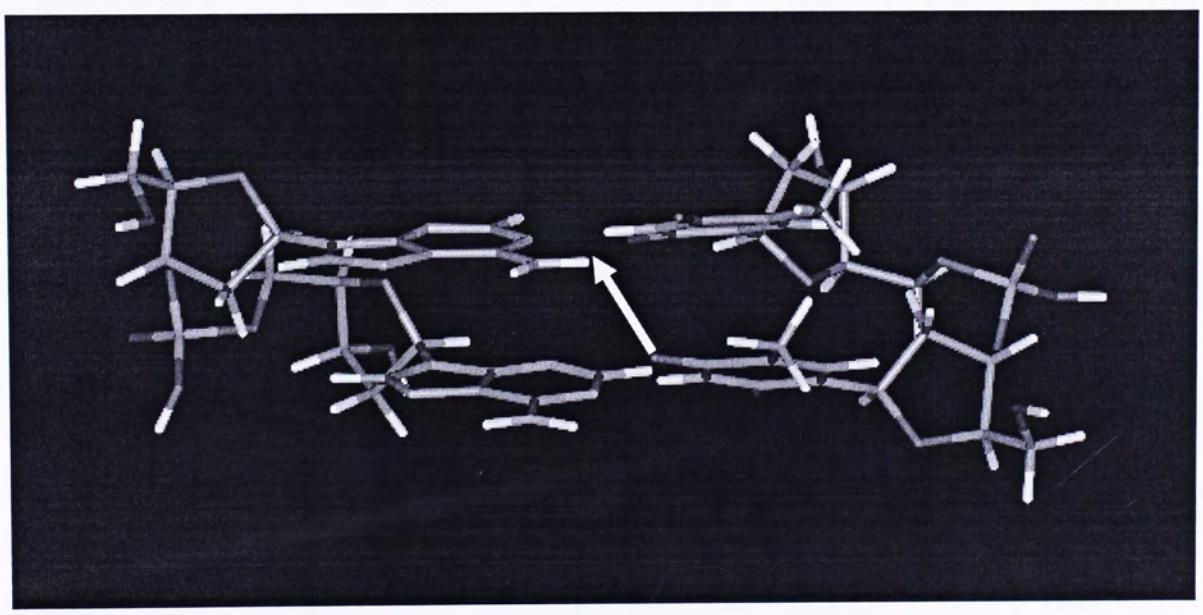


Figure 3.1.1.2 Three-centered H-bond in B Dimer

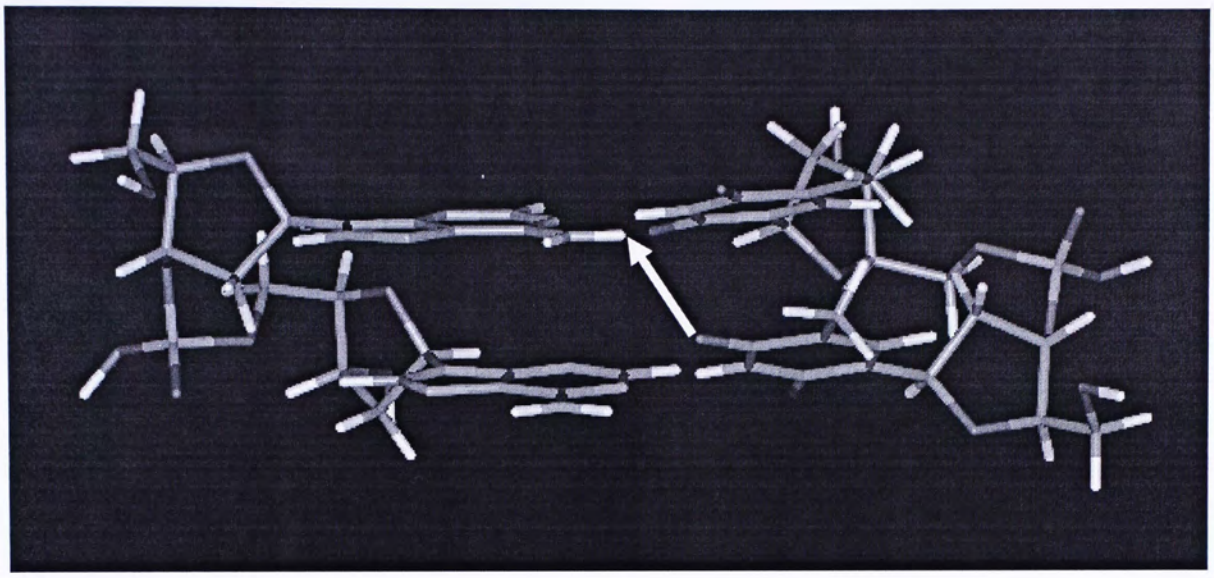


Figure 3.1.1.3 Three-centered H-bond in C Dimer

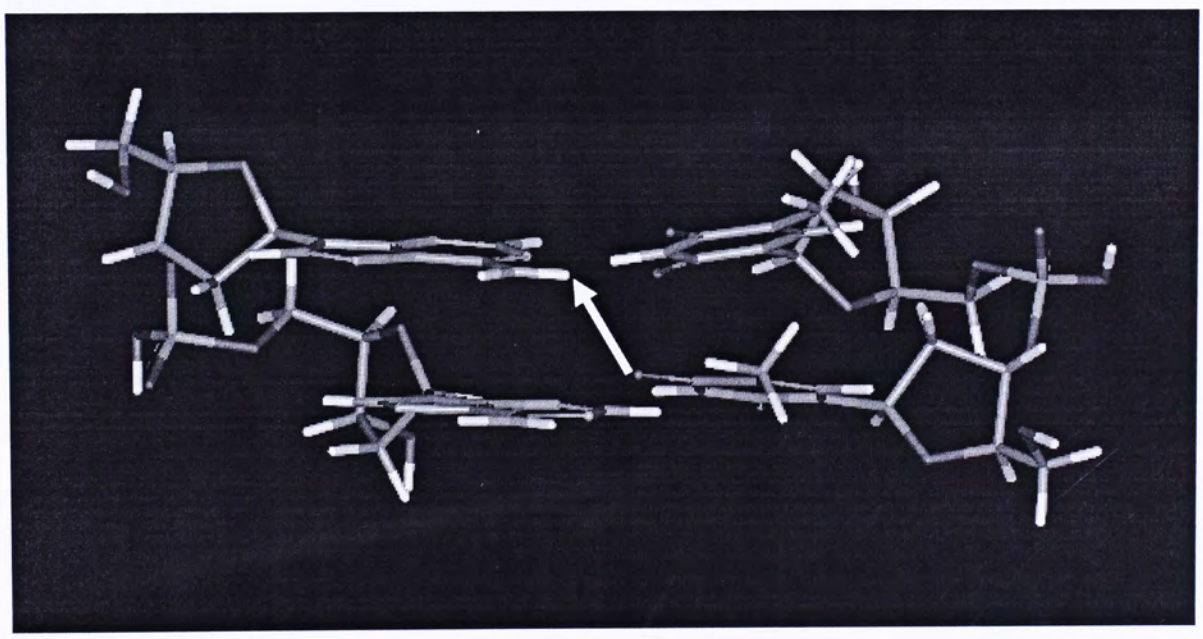


Figure 3.1.1.4 Three-centered H-bond in D Dimer

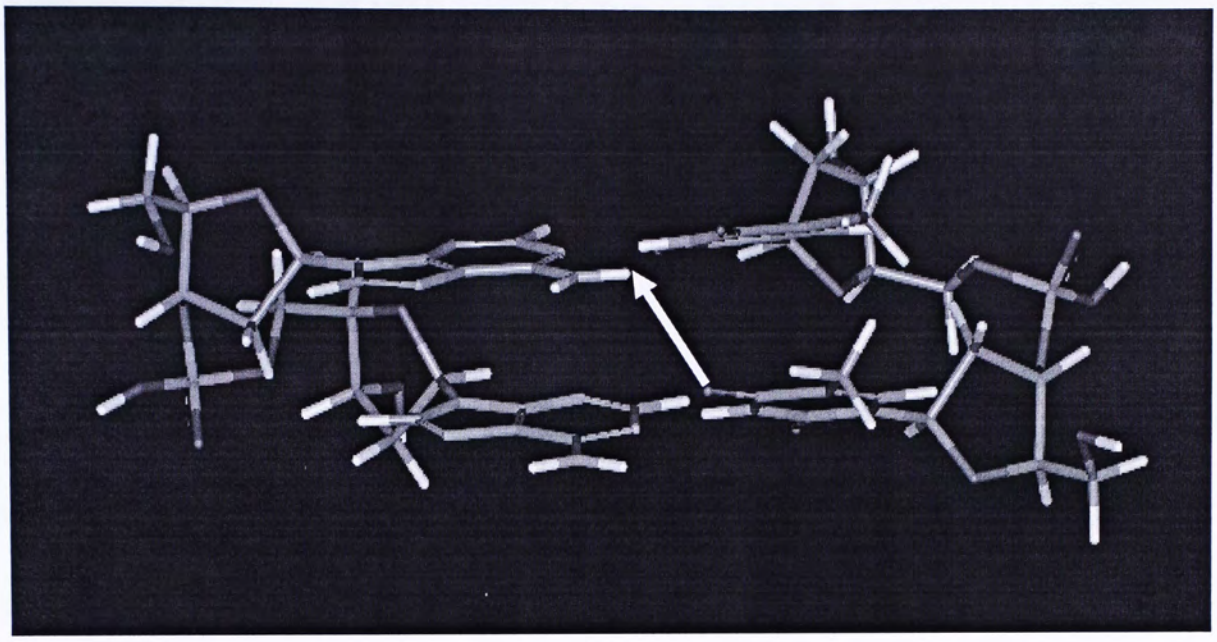


Figure 3.1.1.5 Three-centered H-bond in E Dimer

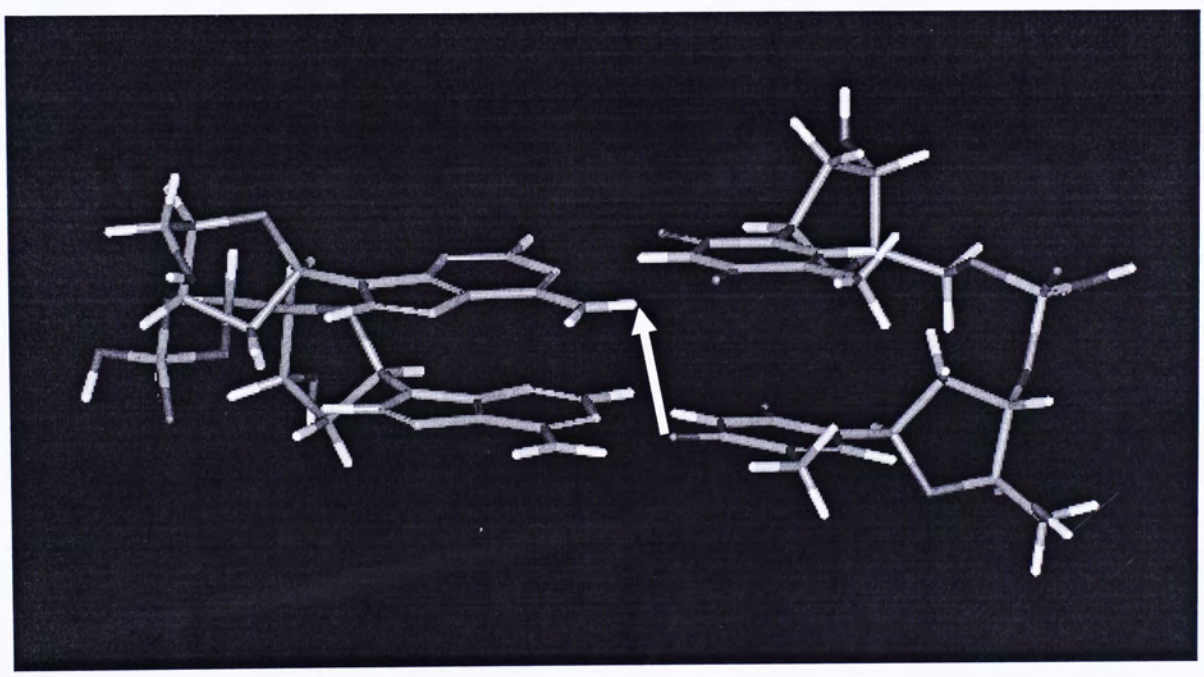


Figure 3.1.1.6 Three-centered H-bond in F Dimer

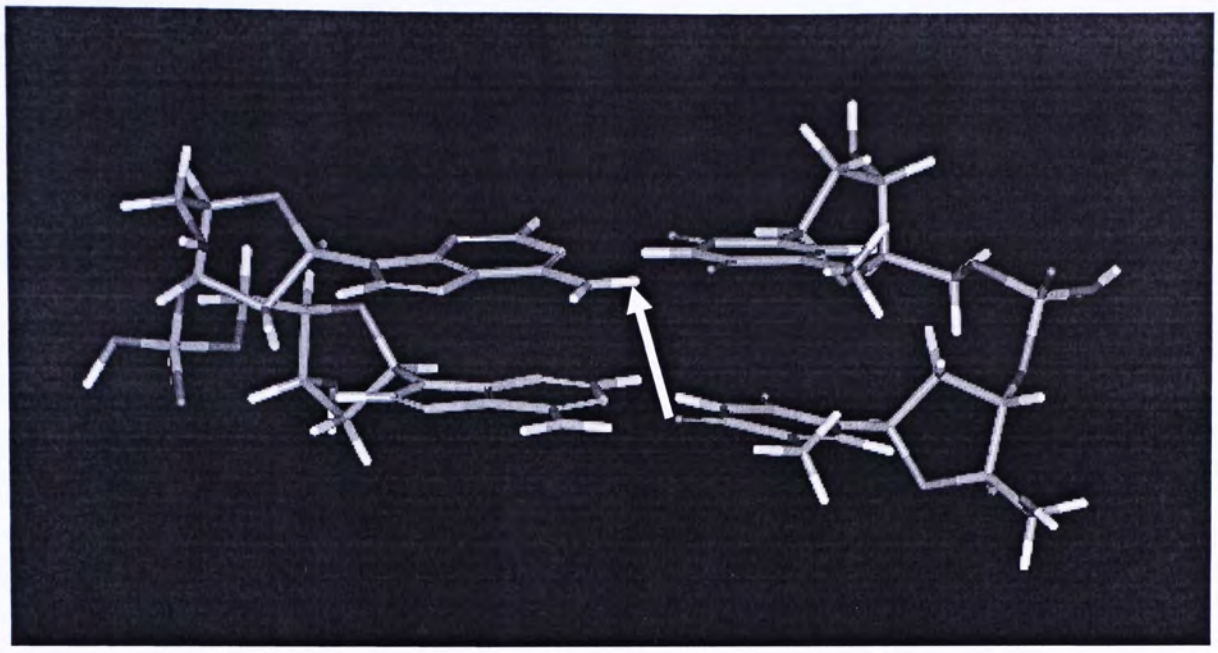


Figure 3.1.1.7 Three-centered H-bond in G Dimer.

Table 3.1.1.1 The geometry information of three-centered and intermolecular two-centered H-bonds

A8A9	Bond distance (Å)			Angle (°)			SAN (°)[g]
dimer	r(A-H)[a]	r(H…B)[b]	r(H…B')[c]	A(AHB)[d]	A(AHB')[e]	A(BHB')[f]	
A	1.01	2.25	2.33	148.6	123.3	87.9	359.8
B	1.01	2.17	2.45	154.5	123.2	82.0	359.7
C	1.01	2.21	2.53	156.1	124.2	79.5	359.6
D	1.01	2.23	2.58	150.5	129.3	80.0	359.8
E	1.01	2.12	2.64	159.0	122.7	77.2	358.9
F	1.01	1.79	2.81	166.4	105.4	87.3	359.1
G	1.01	1.65	2.95	175.6	103.3	81.0	359.9

A is the proton donor

H is the hydrogen atom

B and B' are the proton acceptors

a. r(A-H) denotes the covalent bond distance between A and H.

b. r(H…B) denotes the intermolecular hydrogen bond distance between H and B.

c. r(H…B') denotes the three-centered hydrogen bond distance between H and B'

d. A(AHB) denotes as the angle formed involve A, H and B

e. A(AHB') denotes as the angle formed involve A, H and B'

f. A(BHB') denotes as the angle formed involve B, H and B'

g. SAN is the summation of the bond angles formed around the H atom

It is observed that four of the seven three-centered H-bond distances are shorter than the vdw radius of 2.6 Å for a typical N-H…O bond (3.1), with the largest value of 2.95 Å and 2.33 Å being the smallest. However, from the experimental point of view, all of the seven three-centered H-bonds are suggested to be regarded as three-centered H-bonds which are within the bond distance range of 2.2 – 3.2 Å (3.2). The average bond length of the three-centered H-bonds in the AA tract structure, 2.8 Å, is similar to those reported by Nelson *et al.*(3.3). Although all of the seven three-centered H-bonds conform to the geometry requirement to be considered as true H-bonds, it is suggested that the use of the vdw criteria for strong hydrogen bonds is applicable, but not for the weaker bond nature of three-centered H-bonds.

The bond distance of the intermolecular two-centered hydrogen bonds, range from 1.65 Å to 2.25 Å. According to Jeffery (3.2), the intermolecular two centered N-H…O hydrogen bond can be classified as the moderate type with bond length in the proposed range of 1.5 Å to 2.2 Å and bond angle larger than 130°.

There is a geometrical feature of hydrogen bonds that is related to the three-centered hydrogen bond angles, the Parthasarthy SAN rule (3.4) which can be obtained by adding up the three angles around the hydrogen atom. Figure 3.1.1.9 displays the bond angles formed between A(A8(N6)-A8(H6a)…T16(O4)), A(A8(N6)-A8(H6a)…T15(O4)) and A(T16(O4)…A8(H6a)…T15(O4)) for each A8A9 dimers. It is observed that the equilibrium geometries of the A8A9 dimers obtained from different sequences are not significantly different from each other. All the seven three-centered H-bonds have SAN in between the range of 350°–360°, which confirms that they lie almost in the same plane. Therefore, they conform to the

SAN rule. This is mainly due to the result of the significant propeller twist in the AT pairs. In fact, it has been demonstrated that the large propeller twist considerably improves the three-centered H-bond distance and angle between the base-pair (3.3). It is believed that the tendency towards planarity for the three-centered H-bond is analogous to the tendency towards linearity for the two-centered hydrogen bond. The intermolecular two-centered hydrogen bond angles A8(N6)-A8(H6a)...T16(O6) in the seven A8A9 studied dimers present average values of 148° to 175°, indicating a trend toward linearity in these bonds. It is interesting to note that the greater the directional character of these intermolecular two-centered hydrogen bond angles, the stronger the three-centered hydrogen bond measured, except for the "D Dimer".

In general, the AHB' angles are larger than 90°. According to Jeffery (3.2), this important deviation from linearity for all three-centered H-bonds indicates that the directionality is weak and so are their interactions. Referring to the bond distances and bond angles tabulated in Table 3.1.1.1, no direct correlation is found between these two structural parameters.

Three-centered H-bonds with the same N-H...O hydrogen bond type are found for each A8A9 dimers. We believe this observation could be attributed to two main reasons. Firstly, owing to the nonplanar nature of the out-lying NH_2 amino groups of the adenine base, the two amino hydrogens deviate from the nucleobase plane in one direction and easily accommodate an orientation that is least steric hindered with the purine ring. In fact the NH_2 groups are pyramidal with sp^3 hybridization on their central nitrogen (3.5-3.7). Such phenomena can be shown by looking at the NMR structure of the adenine base displayed in Figure 3.1.1.8. In addition, the involvement of the amino NH_2 group in the three-centered H-bond interactions is probably

responsible for the large propeller twisting of AA tract which forces the N-H group from the adenine base more closer to the oxygen atom of the diagonal thymine base.

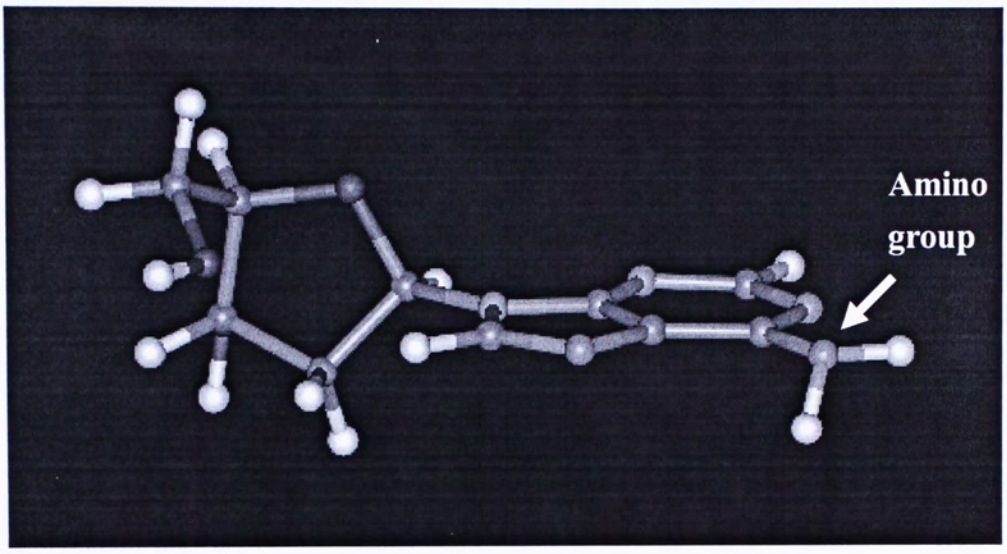


Figure 3.1.1.8 Pyramidal geometry of adenine amino group

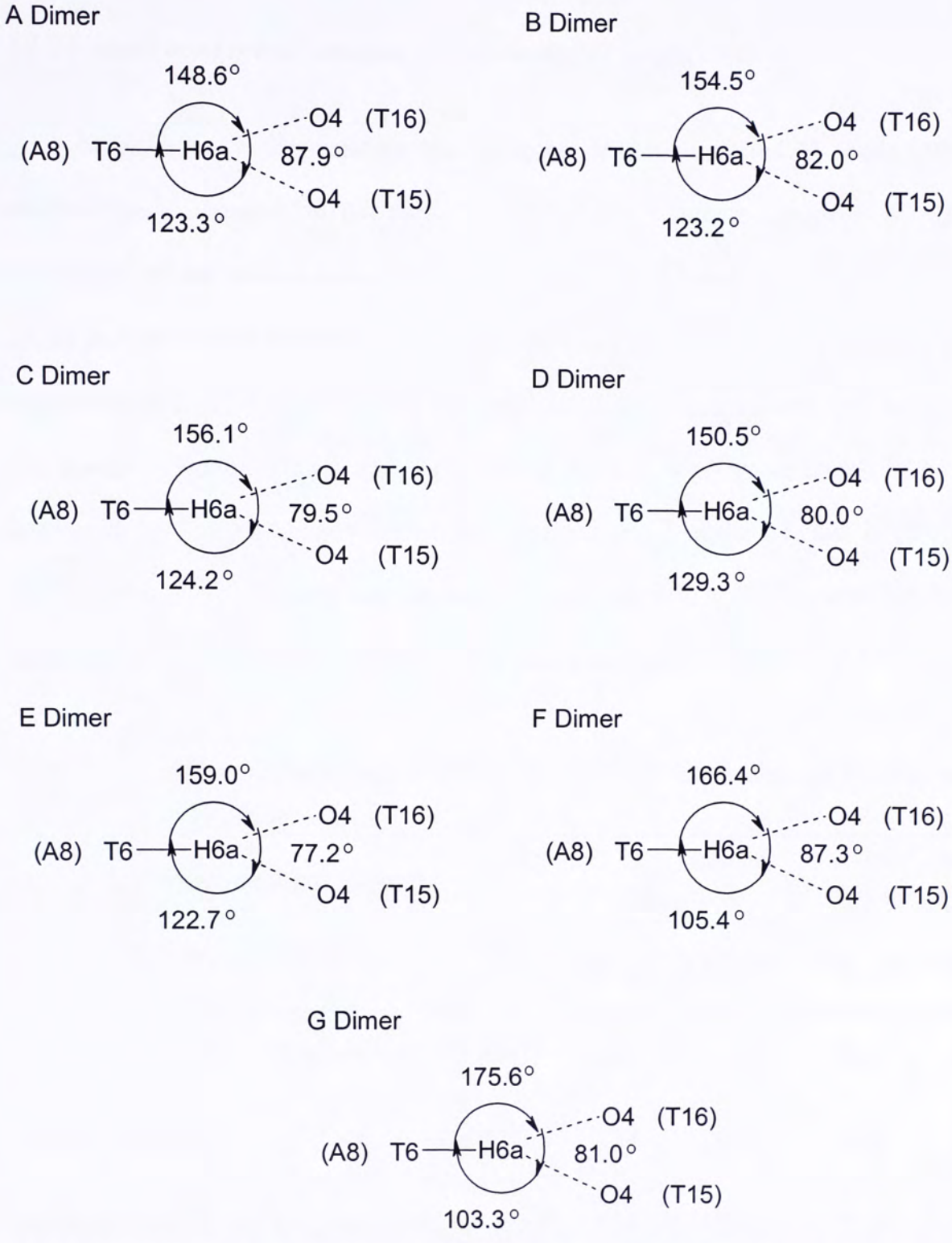


Figure 3.1.1.9 Schematic illustration of the SAN of three-centered hydrogen bonds found in the A8-A9/T16-T15 dimer

3.1.2 Natural bond orbital analysis – Donor-accepter interactions

In order to further analyse the presence of three-centered H-bonds, NBO analysis was conducted on the partially optimized AA dimers. Together with the calculation of the second-order perturbation energy, the orbital interaction can be traced and the values obtained can be used as a way to quantitatively estimate the relative strength of hydrogen bonds. The larger the E(2) value indicates that the more the energy lowering effect. The E(2) which reflects the three-centered H-bond interaction in A8A9 dimers could be obtained by looking at the LP(O4)→ BD*(N6-H6a) delocalization and the sum of each lone pair to BD* interaction. The NBO results calculated with B3PW91/6-311G(d,p) are listed in Table 3.1.2.1.

Table 3.1.2.1 The E(2) information of three-centered and intermolecular two-center H-bonds

A8A9 dimer	Three-centered H-bond			E(2) energy (kcal/mol)		Bond distance (Å)	
	Donor (A-H)	Acceptor (B)	Acceptor (B')	H...B	H...B'	H...B	H...B'
A	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	3.00	1.26	2.25	2.33
B	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	4.24	0.88	2.17	2.45
C	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	3.64	0.62	2.21	2.53
D	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	3.18	0.60	2.23	2.58
E	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	4.96	0.44	2.12	2.64
F	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	16.36	0.10	1.79	2.81
G	A8(N6)-A8(H6a)	T17(O4)	T16(O4)	26.62	0.04	1.65	2.95

To understand the possible correlations between three-centered H-bond energy and the three-centered H-bond geometry, an attempt has been made to quantify the bond strength and determine the feasible existence of three-centered H-bond in DNA dimers. Seven DNA A8A9 dimers which were obtained from two sources of NMR structures are targeted. To investigate the three-centered H-bond on the effect of their bond distance, these two parameters are compared and tabulated (Table 3.1.2.1).

Seven LP(O)→ BD*(N-H) three-centered H-bonding interactions with E(2) values in between 0.04 kcal/mol and 1.26 kcal/mol of A8A9 dimers are summarized (Table 3.1.2.1). It seems that the intermolecular two-centered hydrogen bonding interactions become weaker when simultaneously another three-centered H-bond interaction takes place. In another words, this complementary situation could be attributed to the competition between the formation of two-centered and three-centered hydrogen bonds for the sharing of the same proton donor. Upon strengthening of the two-center hydrogen bonds, antibond of AH aligns closer to the proton acceptor of the two-centered hydrogen bond. Hence, the LP→BD* delocalization for the three-centered H-bonds turn out to be less efficient and less readily formed (3.16).

Data plotted in Figure 3.1.2.2 shows that the shorter the three-centered H-bond distances, the larger the E(2) values of the three-centered H-bonds. A similar trend is observed for the intermolecular two-centered hydrogen bonds, confirming that hydrogen bond strength, is distance dependent. Since E(2) measures the energy lowering effect from the orbitals interaction between the lone pair of proton acceptor and the anti-bonding proton-proton donor orbital (3.8), the shorter distance provides a more feasible orbital overlapping and thus the charge transfer efficiency. This explains

why the intermolecular two-centered hydrogen bonds are generally stronger than the three-centered H-bond because of their short bond distances. Moreover, it is observed that E(2) decreases exponentially as the hydrogen bond distance increases and vanishes gradually as r goes beyond the value of 3.0Å.

A comparison of the trends between the intermolecular two centered hydrogen bonds and the three-centered H-bonds reveals that small differences in the change of hydrogen bond distance in the case of intermolecular hydrogen bonds correspond to fairly substantial changes in their bond strength, but not for the three-centered H-bonds. Hence, this suggests that the strength of intermolecular hydrogen bonds generally posses a larger extent of covalent character in the hydrogen bonding interaction.

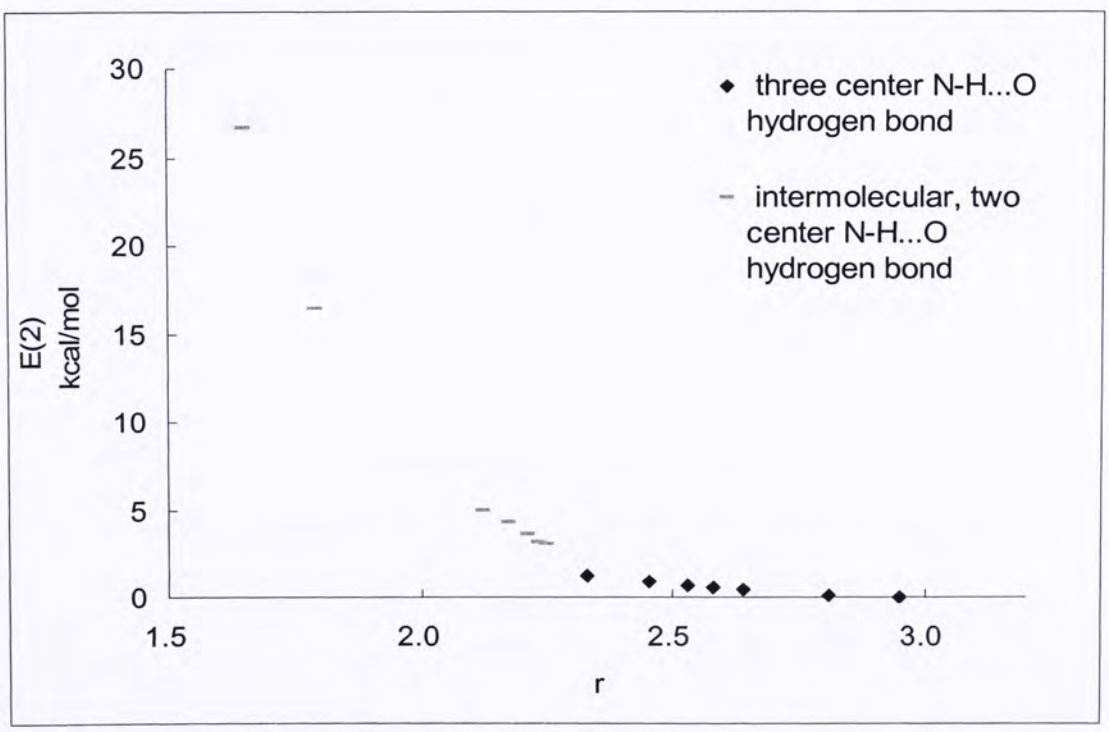


Figure 3.1.2.2 The relationship between E(2) energy and hydrogen bond distance of the intermolecular and three-centered H-bonds of A8A9 dimers optimized at B3PW91 functional

To further examine the correlation between these two energetic and structural parameters a plot of log E(2) versus 1/r is studied as illustrated in Figure 3.1.2.3. On the Y-axis, two separate regions can be discerned at the baseline logE(2)=0. When log E(2) is negative it means that the bond strength is smaller than 1 kcal/mol. The strong intermolecular two-centered hydrogen bonds are mainly distributed on the upper positive region and most of the three-centered H-bonds are on the lower negative region. These two groups of data are fitted to two lines, with respective correlation coefficient of 0.9937 and 0.8752, and with an intersecting point representing 1/r = 0.44 $Å^{-1}$. Trend line extrapolation leads to the following approximate asymptotic values. If computational error of 0.06 kcal/mol is anticipated, then it is suggested that three-centered H-bond with 1/r larger than 0.339 $Å^{-1}$ or r smaller than 2.95 Å can be considered as a true three-centered H-bonding interaction. Such observations can tentatively provide a way to determine the existence of three-centered H-bond base on the calculated E(2) values, especially when little experimental data are available. However, it is believed that a separate plot should better be implemented for different type of hydrogen bonds as a few variations on correlation coefficient, different trend line cut-off value are anticipated.

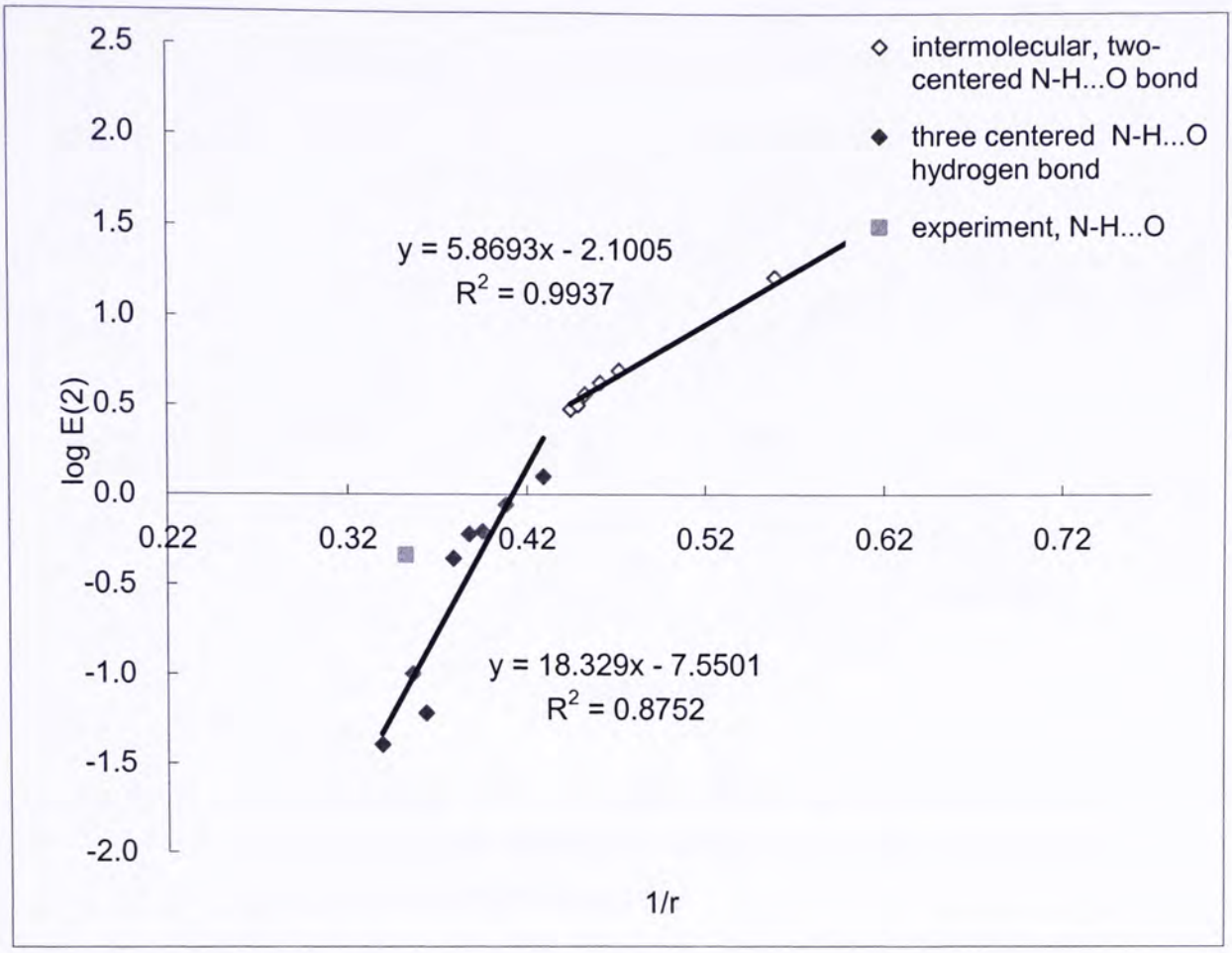


Figure 3.1.2.3 A plot of second-order perturbation theory of E(2) energies of intermolecular and three-centered hydrogen bonding interaction of A8A9 dimers optimized at UB3PW91 functional

Figure 3.1.2.4 shows the relationship between E(2) and hydrogen bond angles of the intermolecular two center hydrogen bonds and the three-centered H-bonds of A8A9 dimers. A trendline is drawn with correlation coefficient of 0.9327. The plot demonstrates that the measured hydrogen bond angles exert a certain effect on hydrogen bond strength. The larger the bond angle towards linearity, the greater the energy lowering effect. However, when considering the bond strength of three-centered H-bond, both bond angle and bond distance parameters should not be consider alone as it is likely the strength of hydrogen bonds do not only depend on the hydrogen bond distance but are also highly directional.

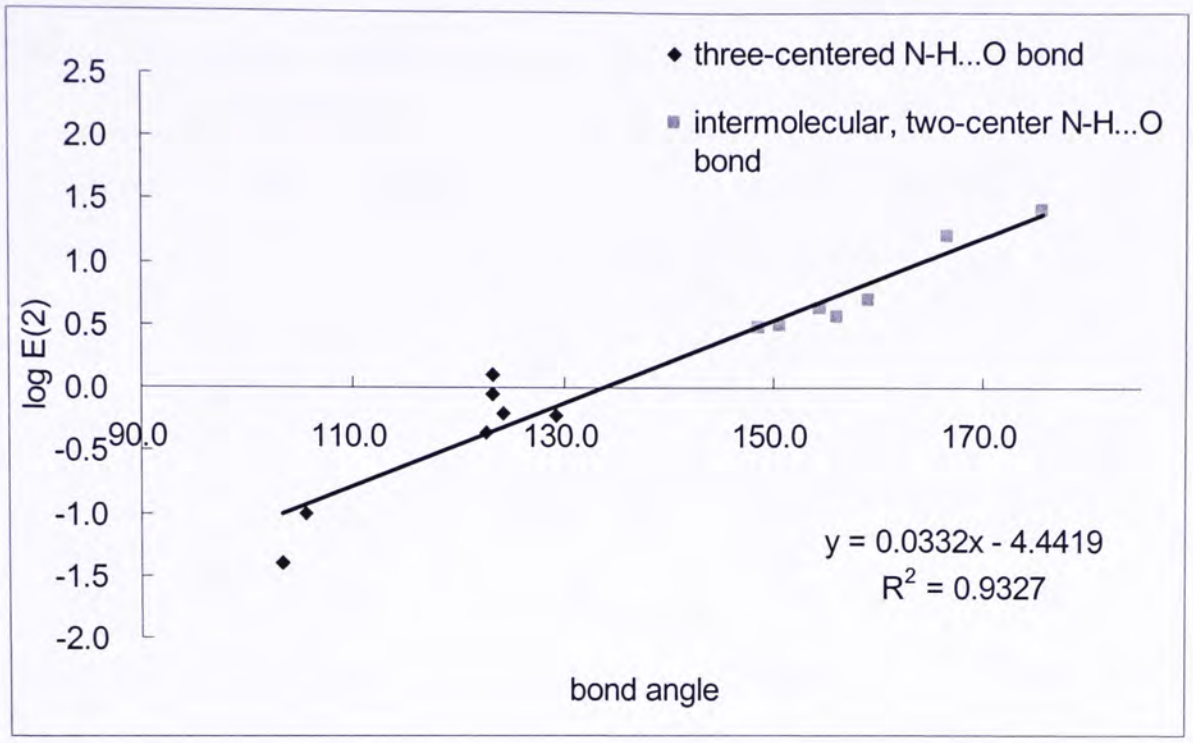


Figure 3.1.2.4 The relationship between log(2) and hydrogen bond angle of intermolecular and three-centered H-bonds of the A8A9 dimers optimized at B3PW91 functional

3.1.3 Spin-Spin coupling across the hydrogen bonds, $^{1h}J_{XH}$

Although three-centered H-bond is not a strong hydrogen bond with full electrostatic character, its covalent character can be shown by looking at the scalar coupling interactions across the hydrogen bridge.

Trans-hydrogen bond J-coupling constant can provide an alternative way to understand the essential role of three-centered hydrogen bonds in biological system which is hardly probed experimentally. Both the three-centered and intermolecular two-center trans-hydrogen bond J-coupling constants ($^{1h}J_{XH}$) between the hydrogen and distant proton acceptor, oxygen, are computed and shown in Table 3.1.3.1.

Table 3.1.3.1 Scalar coupling constants $^{1h}J_{HX}$ of the nuclei between H and X across two base pair.

A8A9 dimer	Scalar coupling constant (Hz)		Bond distance (Å)	
	$^{1h}J_{HB}$ [a]	$^{1h}J_{HB'}$ [b]	r(H...B) [c]	r(H...B') [d]
A	9.60	2.36	2.25	2.33
B	47.76	2.41	2.17	2.45
C	11.56	0.36	2.21	2.53
D	11.04	0.10	2.23	2.58
E	17.20	-0.62	2.12	2.64
F	49.66	-1.80	1.79	2.81
G	73.96	-1.08	1.65	2.95

A is the proton donor

H is the hydrogen atom

B and B' are the proton acceptors

a. $^{1h}J_{HB}$ denotes the trans-hydrogen bond j-coupling constants between H and B

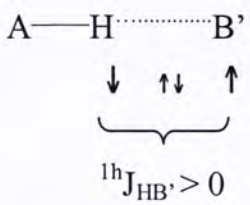
b. $^{1h}J_{HB'}$ denotes the trans-hydrogen bond j-coupling constants between H and B'

c. r(H...B) denotes the intermolecular hydrogen bond distance between H and B.

d. r(H...B') denotes the three-centered hydrogen bond distance between H and B'

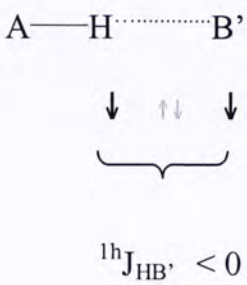
It is observed that all the entries in the third column of Table 3.1.3.1 have a non-zero value. In fact, if three-centered H-bonds possess covalent character, then there would be non-zero J-coupling constant values about the three-centered hydrogen bridges (3.34). However, three out of seven of the three-centered H-bonds have slightly negative J-coupling constant values in between -0.62 to -1.80 Hz, and increase to large positive values as r(H...B') becomes shorter than 2.58 Å. Compared with the positive J-coupling constant reported for the three-centered hydrogen bond,

sizable values are obtained for the two-center hydrogen bonds with r(H...B) becomes smaller than 2.25 Å, J-coupling constant range from 9.60 to 73.96 Hz. Accordingly, Andrew (3.9) reported that the values of the J-coupling constant can either be positive or negative. Pictorial explanations for these findings are given using the following diagram.



Here, ↓ and ↑ represent nuclear spins, and ↑↓ represent the spins of the pair of bonding electrons. Recalling that J-coupling is referred to as spin-spin coupling, two nuclei in a molecule which are connected by one bond, H·····B, can be seen to be coupled. Under the effect of an external magnetic field, the nuclear magnetic moment of the nuclei H can give rise to a magnetic field and force its unpaired electron to reorient. Following the Pauli exclusion principle, the polarization of the electron is then transmitted through the bonding orbital by the electron-electron interactions which in turn polarizes the magnetic nuclei of the distant proton acceptor B. Under this circumstance, a one-bond scalar coupling is then observed between both nuclei. Due to the two possible states 1/2 and -1/2 of the spin system, a negative value of the J-coupling constant may then be observed when the two coupled nuclear spin vectors are parallel with each other. By looking at the diagram above, it is suggested that the nuclear spins of the H and proton acceptor B are parallel to each other hence the negative J-coupling values when r(H...B') goes beyond 2.58 Å. One possibility is due to the long three-centered H-bond distances that provide a poor electron-electron interaction with a relatively poor orbital overlap. Hence, the electron spins can no longer strongly polarize the magnetic nuclei of the distant proton acceptor B', and the

alignment of the nuclear magnetic moments for the H and B' pairs stay parallel, as shown in the following scheme.



It is expected that contributions to the $^{1h}J_{XH}$ will be dominated by the type of atoms involves in the coupling and a uniform sign of J-coupling will be anticipated. However it seems that this is not the case for the three-centered H-bonds. The plot in figure 3.1.2.5 shows that as H...B distance increases, the $^{1h}J_{XH}$ decreases but the variation is nonlinear. The same phenomenon can be observed with the two-centered H-bond case. In fact, the distance between the coupling atoms is a dominant factor in determining the value of the coupling constant and explains the reason why the $^{1h}J_{XH}$ for the three-centered H-bonds are generally smaller in magnitude (3.10). It is also observed that a small variation of intermolecular hydrogen bond distance causes a comparatively large change in the spin-spin coupling constant value which correlates well with the results discussed in section 3.2

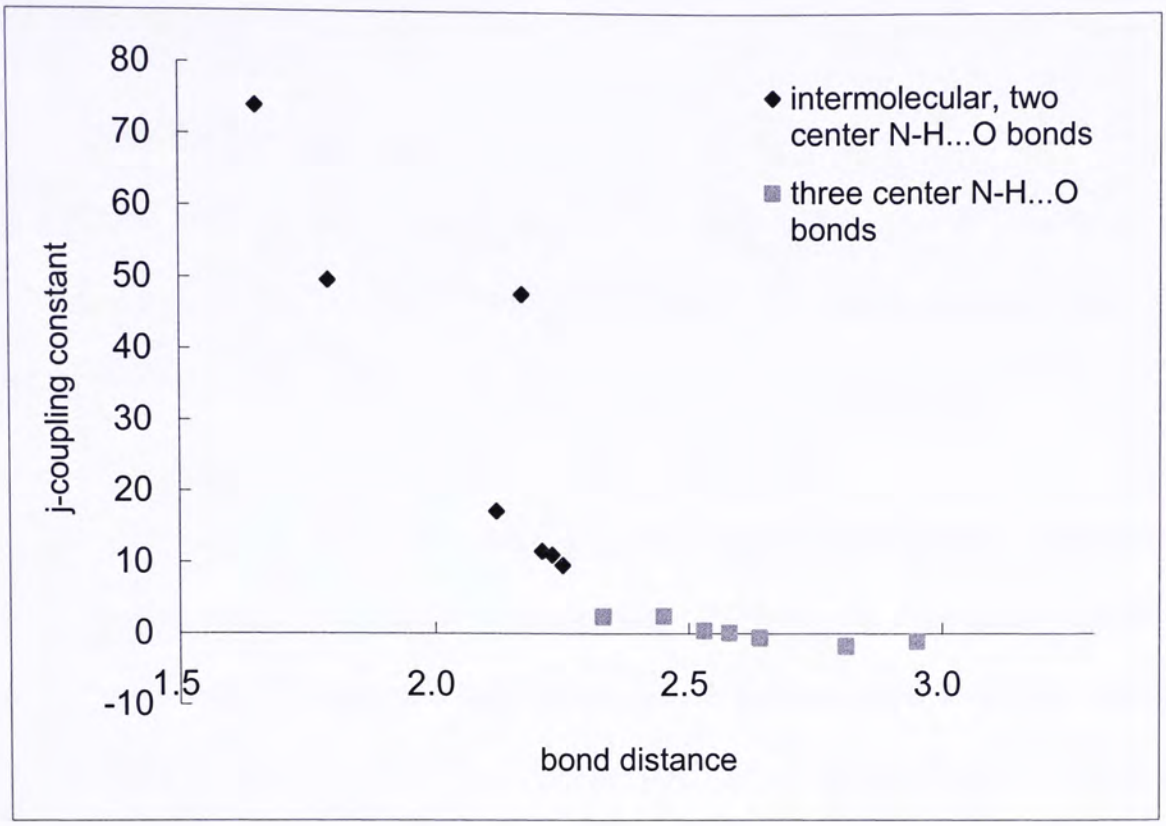


Figure 3.1.2.5 The relationship between j-coupling constant across the hydrogen bonds and hydrogen bond distance of intermolecular and three-centered H-bonds

Together with the results displayed in Table 3.1.1.1 a positive correlation between $^{1h}J_{XH}$ values and hydrogen bond angles is observed, except for the "D Dimer". In fact, a recent study done by Scheurer el. al. (3.11) shows that the H-bond scalar couplings reduced significantly by approximately 20 to 30% for a decrease in the hydrogen bond angle to 140°.

3.1.4 Wiberg bond index

From the NBO analyses, the Wiberg bond index (WBI) can be obtained, which is a related measure of MO bond order. As one of the simplest way to measure the bond order of chemical bonds, a standard WBI value 1.0 is assigned as the bond index for a C-C single bond, while it is 2.0 for a C=C double bond (3.12).

According to the NBO analyses, there exist donor-acceptor interactions between the oxygen lone-pair electrons orbital of O4 and the antibonding orbital of N-H group in all of the seven A8A9 dimers, which are consistent with their non-zero WBI values obtained in Table 3.1.4.1. The WBIs of all three-centered H-bonds are smaller than 1.0, ranging from 0.0009 to 0.0052, implying that they have relatively weak donor-acceptor interactions. The data shown in the same table clearly shows a gradual increase of WBI for the three-centered hydrogen bond with decreasing three-centered H-bond distance, except for the "C Dimer". Judging from analyses on the WBIs and the E(2) of the intermolecular A8(N6)-A8(H6a)...T15(O4) bond in "G Dimer", it seems that the bond order is quite small even have the largest E(2) energy value. In fact, unlike the classical bond-order itself, WBI is considered an approximation to the covalent portion bond order and justified as more overlap equals more "bonds" (3.13). Therefore, the trend of the WBI value generally coincides with the trend of the E(2) energies in Table 3.1.2.1 which depends on the degree of orbitals overlapping. Nevertheless, WBI can be used as a complementary index to examine the degree of covalency of the hydrogen bonds in all seven A8A9 dimers.

A positive correlation is obtained between bond order and E(2) energy as demonstrated in Figure 3.1.2.6.

The relationship between WBI bond order and E(2) is illustrated in Figure 3.1.2.6. The results show that the strength of hydrogen bonds is proportional to the bond order of the bond between hydrogen and oxygen proton acceptor.

Table 3.1.4.1 Wiberg bond indexes (WBIs) of the hydrogen bonds between H and X across two base pair.

A8A9 dimer	Three-centered H-bond		Bond index		Bond distance (Å)	
	Donor	Acceptor	H...B	H...B'	r(H...B)	r(H...B')
A	A8(N6)-A8(H6a)	T16(O4)	0.0119	0.0052	2.25	2.33
B	A8(N6)-A8(H6a)	T16(O4)	0.0161	0.0036	2.17	2.45
C	A8(N6)-A8(H6a)	T16(O4)	0.0010	0.0028	2.21	2.53
D	A8(N6)-A8(H6a)	T16(O4)	0.0129	0.0031	2.23	2.58
E	A8(N6)-A8(H6a)	T16(O4)	0.0182	0.0021	2.12	2.64
F	A8(N6)-A8(H6a)	T16(O4)	0.0483	0.0009	1.79	2.81
G	A8(N6)-A8(H6a)	T16(O4)	0.0677	0.0004	1.65	2.95

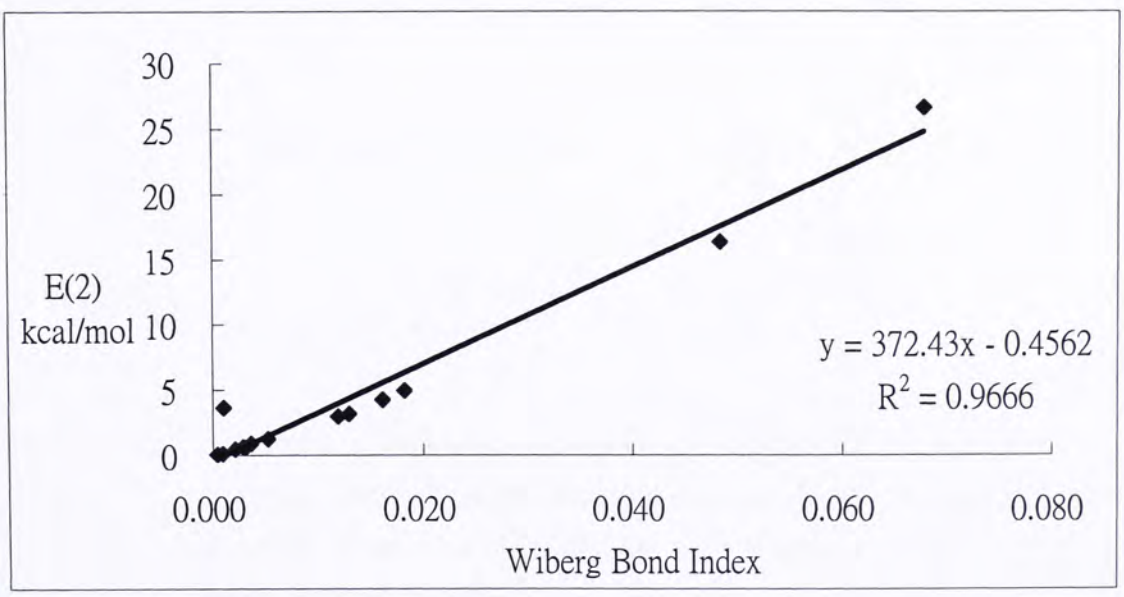


Figure 3.1.2.6 Correlation between E(2) of two center and three-centered component of hydrogen bond and the corresponding Wiberg bond index

3.1.5 Proton transfer - Control of hydrogen bond strengths by a remote proton transfer

It is believed that proton transfer in DNA plays an important role in damaging the integrity of nucleic acid components, and contributes significantly to the cause of DNA mutation via mispairing of complementary bases (3.14). This section focuses on a theoretical study of the effect of hydrogen bond strength with proton transfer reaction taking place within a small DNA oligomer. To shorten the computational time and target on the snap shot result of this reaction, the geometries of the protons involved in the three-centered H-bond formation are manually varied. NBO analysis was then carried out for each r(N-H) adjusted models and the strength of H-bonds are modulated. The schematic representation is shown in Figure 3.1.5.1.

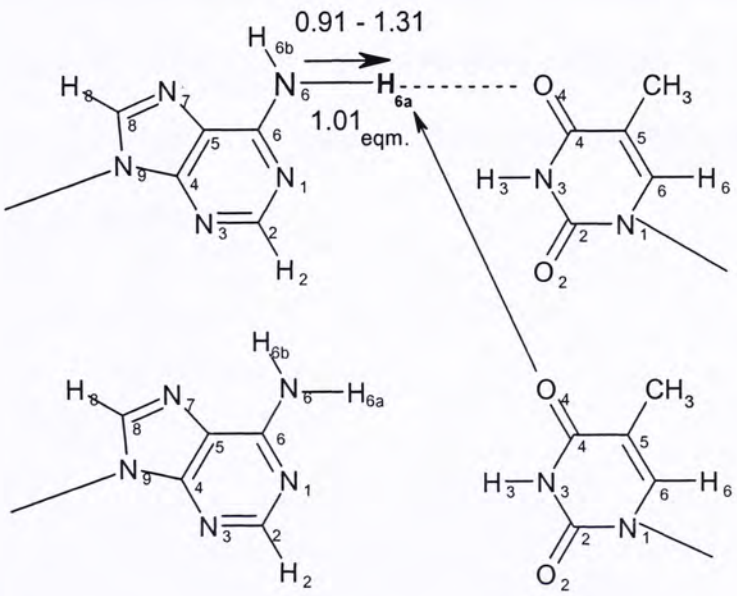


Figure 3.1.5.1 Schematic representation of proton transfer reaction takes place between A-T base pairs. The solid arrow indicates the bond involves in proton transfer reaction and the dash arrow represents the three-centered H-bond

Table 3.1.5.1 and Table 3.1.5.2 summarize the E(2) energies of the hydrogen bonds found in the A7A8A9 trimer before and after intermolecular proton transfer. Also shown in the brankets are the corresponding hydrogen bond distances. The E(2) values (Table 3.1.5.1 and 3.1.5.2) of intermolecular hydrogen bonds rise significantly when the protons A7(H6a) and A8(H6a) were transfered toward the distant proton acceptor oxygen. For example, while the E(2) of the hydrogen bond A7(N6)-A7(H6a)···T18(O4) is 11.12 kcal/mol at equilibrium (bond length = 1.89Å), it changed considerably to 43.55 kcal/mol after extending the bond length of A7(N6)-A7(H6a) to 1.31Å. The corresponding increment for the A7(N6)-A7(H6a)··· T18(O4) hydrogen bond energy approaches 391%. A similar trend is observed for the case of transferring the A8(H6a) (Table 3.1.5.2), but with about 543% enhancement. The large increase in the hydrogen bond energies in both cases confirm the dominant role of the hydrogen bond distance. Hydrogen bond distances go from a value of 1.98Å to 1.61Å and 2.27Å to 1.88Å when A7(H6a) and A8(H6a) respectively move toward the oxygen, the acceptor.

In view of the three-centered hydrogen bonds, a similar trend can also be observed. Due to the non-planar nature of the three-centered H-bond, bond lengths of three-centered H-bonds with hydrogen atoms involving in proton transfer would also be altered. As expected, bond strength of T17(O4)···A7(N6)-A7(H6a) in Table 3.1.5.1 and T16(O4)···A8(N6)-A8(H6a) in Table 3.1.5.2 are not constant. This is because E(2) estimation involves the transfer of electron density from lone pairs of oxygen atom in the proton acceptor, nO, to the antibonding orbital of the N-H bond in the proton donor, σ^*(N-H). As discussed earlier on the relationship between E(2) energy and bond distance, the longer the hydrogen bond distance, the smaller the E(2) energy of the hydrogen bond. It is therefore suggested that shorter distance can provide a better way for orbital interaction as well as overlapping between the orbitals of O atoms and H atoms.

It is anticipated that the alteration of hydrogen bond strength is mostly based on a function of the hydrogen bond lengths. For this reason, an attempt has been made to modulate the strength of those hydrogen bonds with constant bond distances when a distant proton transfer reaction occurs. Referring to Table 3.1.5.1 and Table 3.1.5.2, it is worthwhile to notice that hydrogen bond strength can be varied even keeping its hydrogen bond distance constant. The E(2) energies of intermolecular hydrogen bonds A8(N6)-A8(H6a)···T17(O4) and A8(N1)···T17(H3)-T17(N3) vary inversely in response to the A7(H6a) proton transfer. There is a 0.1 kcal/mol net E(2) decrement for the A8(N6)-A8(H6a)···T17(O4) hydrogen bond, and 0.34 kcal/mol net increment for the second one. It seems that as A7(H6a) is about to depart from the A7(N6) atom, at the same time, the strength of the two hydrogen bonds formed on the other adenine-thymine base pairs will also be affected even without any structural alteration. A similar phenomenon is observed in the case of A8(H6a) proton transfer as

illustrated by the data in Table 3.1.5.2. In view of the E(2) energy changes of the three-centered H-bonds along with the oligomer size, it is suggested that their bond strengths are less likely to be affected, in response to the distance proton transfer reaction.

Table 3.1.5.1 E(2) energy of the hydrogen bonds found in A7A8A9 trimer with A7(H6a) proton transfer.

Hydrogen bonds	E(2) kcal/mol				
	r(A7(N6)-A7(H6a)) Å				
Two-centered H-bonds	0.91	1.01_{eqm}	1.11	1.21	1.31
A7(N6)-A7(H6a)···T18(O4)	6.12(1.98)	11.12(1.89)	17.91(1.79)	28.81(1.70)	43.55(1.61)
A7(N1)···T18(H3)-T18(N3)	30.40(1.71)	30.46(1.71)	30.94(1.71)	31.22(1.71)	31.56(1.71)
A7(C2)-A7(H2)···T18(O2)	0.32(2.70)	0.32(2.70)	0.32(2.70)	0.30(2.70)	0.30(2.70)
A8(N6)-A8(H6a)···T17(O4)	4.28(2.17)	4.24(2.17)	4.22(2.17)	4.20(2.17)	4.18(2.17)
A8(N1)···T17(H3)-T17(N3)	34.34(1.69)	34.26(1.69)	34.48(1.69)	34.56(1.69)	34.68(1.69)
A8(C2)-A8(H2)···T17(O2)	0.84(2.44)	0.84(2.44)	0.84(2.44)	0.84(2.44)	0.84(2.44)
A9(N6)-A9(H6a)···T16(O4)	1.34(2.38)	1.34(2.38)	1.34(2.38)	1.34(2.38)	1.34(2.38)
A9(N1)···T16(H3)-T16(N3)	31.02(1.70)	30.96(1.70)	31.04(1.70)	31.04(1.70)	31.06(1.70)
A9(C2)-A9(H2)···T16(O2)	0.68(2.45)	0.68(2.45)	0.64(2.45)	0.64(2.45)	0.68(2.45
Three-centered H-bonds					
T17(O4)···A7(N6)-A7(H6a)	0.04(3.01)	0.14(2.97)	0.28(2.92)	0.52(2.87)	0.78(2.83)
T16(O4)···A8(N6)-A8(H6a)	0.88(2.45)	0.88(2.45)	0.88(2.45)	0.88(2.45)	0.88(2.45)
A8(O2)···A9(C2)-A9(H2)	0.02(2.72)	0.02(2.72)	0.02(2.72)	0.02(2.72)	0.02(2.72)

Table 3.1.5.2 E(2) energy of the hydrogen bonds found in A7A8A9 trimer with A8(H6a) proton transfer

Hydrogen bonds	E(2) kcal/mol				
	r(A8(N6)-A8(H6a)) Å				
Two-centered H-bonds	0.91	1.01_{eqm}	1.11	1.21	1.31
A7(N6)-A7(H6a)···T18(O4)	11.24(1.98)	11.12(1.89)	11.22(1.79)	11.20(1.70)	11.12(1.61)
A7(N1)···T18(H3)-T18(N3)	30.70(1.71)	30.46(1.71)	30.76(1.71)	31.80(1.71)	31.74(1.71)
A7(C2)-A7(H2)···T18(O2)	0.16(2.70)	0.32(2.70)	0.32(2.70)	0.30(2.70)	0.30(2.70)
A8(N6)-A8(H6a)···T17(O4)	1.90(2.27)	4.24(2.17)	7.90(2.07)	14.20(1.97)	23.02(1.88)
A8(N1)···T17(H3)-T17(N3)	34.20(1.69)	34.26(1.69)	34.58(1.69)	34.78(1.69)	34.92(1.69)
A8(C2)-A8(H2)···T17(O2)	0.84(2.44)	0.84(2.44)	0.84(2.44)	0.84(2.44)	0.82(2.44)
A9(N6)-A9(H6a)···T16(O4)	1.34(2.38)	1.34(2.38)	1.32(2.38)	1.30(2.38)	1.28(2.38)
A9(N1)···T16(H3)-T16(N3)	30.96(1.70)	30.96(1.70)	31.08(1.70)	31.16(1.70)	31.26(1.70)
A9(C2)-A9(H2)···T16(O2)	0.68(2.45)	0.68(2.45)	0.64(2.45)	0.64(2.45)	0.68(2.45)
Three-centered H-bonds					
T17(O4)···A7(N6)-A7(H6a)	0.14(2.97)	0.14(2.97)	0.14(2.97)	0.12(2.97)	0.12(2.97)
T16(O4)···A8(N6)-A8(H6a)	0.40(2.50)	0.88(2.45)	1.56(2.42)	2.56(2.38)	3.60(2.36)
A8(O2)···A9(C2)-A9(H2)	0.02(2.72)	0.02(2.72)	0.02(2.72)	0.02(2.72)	0.02(2.72)

Table 3.1.5.3 E(2) energy of the hydrogen bonds found in A8A9C10 trimer with A8(H6a) proton transfer

Hydrogen bonds	E(2)				
	r(A8(N6)-A8(H6a)) Å				
Two-centered H-bonds	0.91	1.01_{eqm}	1.11	1.21	1.31
A8(N6)-A8(H6a)···T17(O4)	2.16(2.26)	4.24(2.17)	7.78(2.08)	13.20(1.99)	21.08(1.90)
A8(N1)···T17(H3)-T17(N3)	33.00(1.69)	33.14(1.69)	33.30(1.69)	33.48(1.69)	33.70(1.69)
A8(C2)-A8(H2)···T17(O2)	0.82(2.44)	0.82(2.44)	0.82(2.44)	0.82(2.44)	0.82(2.44)
A9(N6)-A9(H6a)···T16(O4)	1.68(2.38)	1.66(2.38)	1.64(2.38)	1.60(2.38)	1.60(2.38)
A9(N1)···T16(H3)-T16(N3)	31.00(1.69)	30.94(1.69)	31.12(1.69)	31.20(1.69)	31.28(1.69)
A9(C2)-A9(H2)···T16(O2)	0.68(2.45)	0.68(2.45)	0.68(2.45)	0.68(2.45)	0.68(2.45)
C10(N4)-C10(H4a)···G15(O6)	3.96(2.02)	3.92(2.02)	3.94(2.02)	3.92(2.02)	3.92(2.02)
C10(N3)···G15(H1)-G15(N1)	23.95(1.73)	23.97(1.73)	23.97(1.73)	23.99(1.73)	24.01(1.73)
C10(O2···G15(H2a)-G15(N2)	29.01(1.61)	29.06(1.61)	29.04(1.61)	29.05(1.61)	29.07(1.61)
Three-centered H-bonds					
T16(O4)···A8(N6)-A8(H6a)	0.94(2.51)	0.88(2.45)	1.74(2.40)	2.92(2.35)	4.48(2.30)
G15(O6)···A9(N6)-A9(H6a)	0.92(2.12)	0.92(2.12)	0.92(2.12)	0.92(2.12)	0.92(2.12)
A9(N1)···G15(H1)-G15(N1)	0.04(3.16)	0.04(3.16)	0.04(3.16)	0.04(3.16)	0.04(3.16)

Table 3.1.5.4 E(2) energy of the hydrogen bonds found in A8A9C10 trimer with A8(H6a) proton transfer

Hydrogen bonds	E(2)				
	r(A9(N6)-A9(H6a)) Å				
Two-centered H-bonds	0.91	1.01_{eqm}	1.11	1.21	1.31
A8(N6)-A8(H6a)···T17(O4)	4.24(2.17)	4.24(2.17)	4.24(2.17)	4.24(2.17)	4.24(2.17)
A8(N1)···T17(H3)-T17(N3)	33.16(1.69)	33.14(1.69)	33.14(1.69)	33.16(1.69)	33.32(1.69)
A8(C2)-A8(H2)···T17(O2)	0.82(2.44)	0.82(2.44)	0.82(2.44)	0.82(2.44)	0.82(2.44)
A9(N6)-A9(H6a)···T16(O4)	0.72(2.45)	1.66(2.38)	3.12(2.31)	5.36(2.24)	8.44(2.18)
A9(N1)···T16(H3)-T16(N3)	30.98(1.69)	30.94(1.69)	31.20(1.69)	31.34(1.69)	31.52(1.69)
A9(C2)-A9(H2)···T16(O2)	0.70(2.45)	0.68(2.45)	0.67(2.45)	0.66(2.45)	0.66(2.45)
C10(N4)-C10(H4a)···G15(O6)	4.00(2.02)	3.92(2.02)	3.88(2.02)	3.80(2.02)	3.72(2.02)
C10(N3)···G15(H1)-G15(N1)	23.95(1.73)	23.97(1.73)	24.01(1.73)	24.09(1.73)	24.17(1.73)
C10(O2)···G15(H2a)-G15(N2)	28.98(1.61)	29.06(1.61)	29.04(1.61)	29.12(1.61)	29.16(1.61)
Three-centered H-bonds					
T16(O4)···A8(N6)-A8(H6a)	0.92(2.45)	0.88(2.45)	0.86(2.45)	0.84(2.45)	0.84(2.45)
G15(O6)···A9(N6)-A9(H6a)	0.20(2.17)	0.92(2.12)	1.80(2.09)	3.06(2.05)	4.60(2.02)
A9(N1)···G15(H1)-G15(N1)	0.04(3.16)	0.04(3.16)	0.04(3.16)	0.04(3.16)	0.04(3.16)

To determine if there is no remote control effect for the three-centered H-bond, a A8A9C10 trimer model with the largest three-centered H-bond E(2) energy is studied. Similar to the case discussed earlier on A7A8A9 proton transfer, the geometry of the protons involved in the formation of three-centered H-bonds is varied. Therefore, geometries of A8(H6a) and A9(H6a) are controlled and the results of E(2) energy change are given in Table 3.1.5.3 and Table 3.1.5.4.

Looking at the data in Table 3.1.5.3, it is observed when H6a of adenine base moves toward the proton acceptor oxygen atom. The bond strengths of both N-H…O and N-H…N bonds on the next adenine base were varied. Again, the resultant change in E(2) for the N-H…N bond is greater than that for the N-H…O bond. The E(2) rises from 31.00kcal/mol to 31.28kcal/mol as bond length of A8(N6)-A8(H6a) elongate from 0.91 Å to 1.31 Å, while there is only 0.08kca/mol net E(2) enhancement for the N-H…O bond. In the case of A9(H6a) proton transfer in Table 3.1.5.4, despite the fact that the base cytosine is located below the point in which proton transfer is taken place, the E(2) of both N-H…O and N-H…N bonds are not remain the same. The E(2) change in opposite directions for these two bonds, but the net E(2) change for the N-H…O bond is +0.28kcal/mol which is much larger than that aforementioned in the adenine base, while it is +0.22kcal/mol for the N-H…N bond in the cytosine base. Unlike A-T base pair, all of the three intermolecular hydrogen bonds' strength in C-G base pair are not constant when the geometry of a distance proton is modified. Therefore, this suggests that the strength of intermolecular hydrogen bonds in C-G base pair are quite sensitive towards its neighboring geometry change even though it seems that there is no direct communication between the two base pairs. Based on this finding, one can elucidate the reason for the failure of DNA integrity when there is an abnormal proton transfer happening at a distant point. Therefore, this leads to the

abnormal weakening of intermolecular hydrogen bond.

It is interesting to find that all of the N-H...N hydrogen bonds in both A-T and C-G base pairs are strengthened wherever proton transfer reaction takes place. This reveals a tendency for N-H…N bonds to play an important role in keeping the adenine base pairs with thymine base and ensure that adenine and thymine are always linked together.

In table 3.1.5.4, an unexpected bond energy change is revealed while considering the proton transfer effect on three-centered H-bond. As A9(H6a) moves toward its proton acceptor oxygen, the bond strength of A9(N6)-A9(H6a)···T16(O4) enhances more than four fold from 1.66kcal/mol at equilibrium state to 8.44kcal/mol when the bond length of A9(N6)-A9(H6a) extends to 1.31 Å. At the same time, the strength of three-centered H-bond T16(O4)···A8(N6)-A8(H6a) weaken fairly from 0.92kca/mol to 0.84kcal/mol with the same three-centered H-bond distances. This implies that the reduction of N-H…O three-centered H-bond bond energy is not contributed from the distance factor, but the electronic environment. However, it is likely that charge density analysis is not appropriate for our characterization as the measured charge density may probably a net effect followed by a series of chemical reaction, such as charge delocalization and hyperconjugation. Also, it is hardly possible for us to determine the atomic charge density before and after three-centered H-bond formation by getting rid of the contribution of the two-centered hydrogen bonds. In fact, referring to Table 3.1.5.5 no trend is observed in the Natural Population Analysis (NPA) of each atoms which are involved in the formation of three-centered H-bonds. With the geometry limitation, it reflects the fact that electronic redistribution or charge transfer effect may dominate the shifting of the hydrogen bond strength.

Therefore, it is difficult to determine the effect of electron density on the E(2) variation because of the possible net charge redistribution all over the purine or pyrimidine rings.

Table 3.1.5.5 Natural population analysis of the atoms involve in three-centered H-bond for each A8A9 dimers.

A8A9 dimer	Charge density, NPA		
	N	H	O
A	-0.739	0.436	-0.669
B	-0.737	0.438	-0.659
C	-0.741	0.434	-0.658
D	-0.742	0.436	-0.662
E	-0.740	0.436	-0.652
F	-0.740	0.425	-0.656
G	-0.740	0.423	-0.653

With the simple AAA and AAC trimers on mimicking proton transfer in small DNA oligomers, it is intuitively simple for us to understand the relay effects of hydrogen bridges in DNA and how does it achieves control over hydrogen bonding. An effective remote communication is demonstrated with one proton transfer-triggered in both A7A8A9 and A8A9C10 timer models. Though it does not bring out any significant change on the energetic parameter of the three-centered hydrogen bonds, it is believed that the strength of hydrogen bond can be varied under the effect of distant proton transfer within a small DNA system

According to the definition of E(2) mentioned in chapter two, it is possible for us to find the non-conventional C-H...O hydrogen bonding interaction in our studied systems. In fact, more than one C-H...O interaction is found in our AAA trimer

models. However, the SAN measured deviated much more from 360 degree and failed to fulfill the SAN rule. Consequently, no detailed discussion will be made from the results obtained for this kind of interaction in this study. Still, it is interested to find that in each proton transfer cases, bond strengths of C-H…O bond remain the same. The C-H…O bonds in A-T base pairs are therefore suggested to be the most inert ones and less likely to be influenced wherever proton transfer reaction takes place.

The E(2) parameter in Table 3.1.5.1 and Table 3.1.5.2 shows that hydrogen bond energies found in A-T base pairs follow the descending order of

$$\text{N-H…N} > \text{N-H…O} >> \text{C-H…O}$$

Hydrogen bonds formed between N-H group and N are of considerable strength when compared with other types of hydrogen bonds studied. The bond length difference can account for the relative strength of these three intermolecular hydrogen bonds. In turn the nature of the proton acceptor atoms plays an important role in determining the interaction length. It is suggested that hydrogen bond length becomes shorter when the acceptor atom bears a larger partial negative charge. In fact, NPA results summarized in Table 3.1.5.5 are in line with this observation. Moreover, nitrogen atoms with more s character can spatially provide a better orbital interaction with the H atom. Hence, the directional character, the bond length and the electronic nature of the hydrogen bonds are probably the features that allow one to elucidate its strength.

3.1.6 Cooperative character of hydrogen bonds in short DNA oligomers

To investigate the cooperatative effect with $(AA)_n$ chain length, the sum of the hydrogen bond energies of three of the AA dimer including A6A7, A7A8 and A8A9 in different $(AA)_n$ oligomers are studied. Defined as "E(2) sum" in this study. For instance, in A6A7 dimer, a value of E(2) sum can be obtained by adding up the six intermolecular hydrogen bonds together. Tables 3.1.5.6-3.1.5.8 present the UB3PW91/6-311G** values of the $E(2)_{n\rightarrow\sigma*}$ delocalization energies along the $(AA)_n$ chains. To ensure the change in E(2) does not come from the distance variation factor, all of the bond distances are kept constant.

The E(2) difference of each DNA oligomers is used as a way to measure the degree of cooperativity. Comparison of E(2) values associated with hydrogen bond formation for AA steps in dimer, trimer, tetramer and pentamer units demonstrate that a small cooperative effect takes place upon increasing the length of the oligomers. The E(2) sum of A8A9 dimer enhances slightly from 80.48kcal/mol in its dimer unit to 83.04kcal/mol computed for the pentamer unit. Similar trends are demonstrated for the cases of A7A8 dimer and A6A7dimer in Table 3.1.5.6 and Table 3.1.5.7 respectively, upon increasing the oligomer sizes. Increase stabilization of hydrogen bond interaction is manifested on the increase of E(2) value when compare with that of the dimer unit. The most pronounced hydrogen bond cooperative effect is observed for the case of A6A7 dimer (Table 3.1.5.6), with 9.35kcal/mol (10%) net E(2) strengths from 86.94kcal/mol (in dimer unit) to 96.29kcal/mol (in pentamer unit). This indicates that a positive cooperative effect does exist along the chain length. Therefore, it is suggested that donor-acceptor interactions in (A6A7)n, $(A7A8)_n$ and $(A8A9)_n$ oligomers are cooperated. This phenomenon which is well studied by

Weinhold cannot be explained by the electrostatic nature of hydrogen bonding interaction (3.15), but the charge transfer interaction. Positive cooperativity emerges when a small quantity of charge is transferred between A8A9 and A7A8 in the A8A9→ A7A8 sense through the formation of three-centered H-bonds. A8A9 becomes slightly anionic character $(A8A9)^{\delta-}$ and better donor, while A7A8 becomes slightly cationic in character $(A7A8)^{\delta+}$. Consequently, the A8A9→ A7A8 interaction can be improved in a concerted way. A more efficient A8A9→ A7A8→ A6A7 interaction is anticipated and leads to the largest ΔE(2) sum recorded for the A6A7 dimer. As a way to assist charges move from one place to another, three-centered H-bonds should play an important role and act as a bridgehead to allow this association to happen. All bond distance are kept constant meaning that the change in E(2) does not come from the distance variation factor. However, it is expected that the E(2) sum will not expand in an infinite way instead, the hydrogen bond cooperativity effect seems to initially increase with the oligomer and will approach to an asymptotic value .

Table 3.1.5.6 Intermolecular and three-centered hydrogen bond energies of A6A7 dimer in each size of DNA oligomers

Hydrogen bonds	E(2) energy				r(H...B)
A-H...B	dimer[a]	trimer[b]	tetramer[c]	pentamer[d]	
A5-T20 base pair					
N6-H6a···O4				14.26	1.82
N1···H3-N3				29.60	1.72
C2-H2···O2				0.24	2.72
A6-T19 base pair					
N6-H6a···O4	*22.14*	*21.08*	*22.23*	*22.21*	1.67
N1···H3-N3	*29.06*	*28.80*	*29.46*	*30.26*	1.73
C2-H2···O2	*0.08*	*0.08*	*0.08*	*0.08*	2.98
A7-T18 base pair					
N6-H6a···O4	*10.09*	*11.06*	*11.10*	*11.06*	1.89
N1···H3-N3	*31.30*	*31.74*	*32.12*	*32.26*	1.71
C2-H2···O2	*0.36*	*0.08*	*0.34*	*0.34*	2.70
A8-T17 base pair					
N6-H6a···O4		4.06	4.20	4.14	2.17
N1···H3-N3		33.50	34.30	34.26	1.69
C2-H2···O2		0.88	0.84	0.84	2.44
A9-T16 base pair					
N6-H6a···O4			1.34	1.34	2.38
N1···H3-N3			31.20	31.40	1.70
C2-H2···O2			0.68	0.68	2.45
E(2) sum	*86.94*	*93.62*	*95.33*	*96.29*	
A-H...B'					
A6(N6)-A6(H6a)...T18(O4)	0.06	0.06	0.06	0.06	3.05

a. E(2) energies of the hydrogen bonds calculated in A6A7 dimer.

b. E(2) energies of the hydrogen bonds calculated in A6A7A8 dimer.

c. E(2) energies of the hydrogen bonds calculated in A6A7A8A9 dimer.

d. E(2) energies of the hydrogen bonds calculated in A5A6A7A8A9 dimer

Table 3.1.5.7 Intermolecular and three-centered hydrogen bond energies of A7A8 dimer in each size of DNA oligomers

Hydrogen bonds	E(2) energy				r(H...B)
A-H...B	dimer[a]	trimer[b]	tetramer[c]	pentamer[d]	
A5-T20 base pair					
N6-H6a···O4				14.26	1.82
N1···H3-N3				29.60	1.72
C2-H2···O2				0.24	2.72
A6-T19 base pair					
N6-H6a···O4			22.23	22.21	1.67
N1···H3-N3			29.46	30.26	1.73
C2-H2···O2			0.08	0.08	2.98
A7-T18 base pair					
N6-H6a···O4	*11.08*	*11.12*	*11.10*	*11.06*	1.89
N1···H3-N3	*30.66*	*30.46*	*32.12*	*32.26*	1.71
C2-H2···O2	*0.32*	*0.29*	*0.34*	*0.34*	2.70
A8-T17 base pair					
N6-H6a···O4	*4.08*	*4.24*	*4.20*	*4.14*	2.17
N1···H3-N3	*33.34*	*34.26*	*34.30*	*34.26*	1.69
C2-H2···O2	*0.86*	*0.84*	*0.84*	*0.84*	2.44
A9-T16 base pair					
N6-H6a···O4		1.34	1.34	1.34	2.38
N1···H3-N3		30.96	31.20	31.40	1.70
C2-H2···O2		0.68	0.68	0.68	2.45
E(2) sum	80.48	81.38	83.04	83.04	
A-H...B'					
A7(N6)-A7(H6a)···T17(O4)	0.14	0.14	0.14	0.14	0.96

a. E(2) energies of the hydrogen bonds calculated in A6A7 dimer.

b. E(2) energies of the hydrogen bonds calculated in A6A7A8 dimer.

c. E(2) energies of the hydrogen bonds calculated in A6A7A8A9 dimer.

d. E(2) energies of the hydrogen bonds calculated in A5A6A7A8A9 dimer

Table 3.1.5.8 Intermolecular and three-centered hydrogen bond energies of A8A9 dimer in each size of DNA oligomers

Hydrogen bonds	E(2) energy				r(H…B)
A-H…B	dimer[a]	trimer[b]	tetramer[c]	pentamer[d]	
A5-T20 base pair					
N6-H6a···O4				14.26	1.82
N1···H3-N3				29.60	1.72
C2-H2···O2				0.24	2.72
A6-T19 base pair					
N6-H6a···O4			22.23	22.21	1.67
N1···H3-N3			29.46	30.26	1.73
C2-H2···O2			0.08	0.08	2.98
A7-T18 base pair					
N6-H6a···O4		11.12	11.10	11.06	1.89
N1···H3-N3		30.46	32.12	32.26	1.71
C2-H2···O2		0.29	0.34	0.34	2.70
A8-T17 base pair					
N6-H6a···O4	*4.24*	*4.24*	*4.20*	*4.14*	2.17
N1···H3-N3	*33.26*	*34.26*	*34.30*	*34.26*	1.69
C2-H2···O2	*0.82*	*0.84*	*0.84*	*0.84*	2.44
A9-T16 base pair					
N6-H6a···O4	*1.36*	*1.34*	*1.34*	*1.34*	2.38
N1···H3-N3	*31.24*	*30.96*	*31.20*	*31.40*	1.70
C2-H2···O2	*0.68*	*0.68*	*0.68*	*0.68*	2.45
E(2) sum	72.48	73.20	73.44	73.56	
A-H…B'					
A8(N6)-A8(H6a)···T16(O4)	0.88	0.88	0.88	0.90	2.45

a. E(2) energies of the hydrogen bonds calculated in A6A7 dimer.

b. E(2) energies of the hydrogen bonds calculated in A6A7A8 dimer.

c. E(2) energies of the hydrogen bonds calculated in A6A7A8A9 dimer.

d. E(2) energies of the hydrogen bonds calculated in A5A6A7A8A9 dimer

For the case of three-centered H-bonds, no significant E(2) variation is observed as one increases the size of the oligomers in all three dimers. It is revealed that no cooperative effect is likely to be present.

CHAPTER FOUR
CONCLUDING REMARKS

In this study, we have employed high level ab initio theory to investigate three-centered hydrogen bonds (TCHBs) in DNA. Following the successful results obtained by Pople in calculating the hydrogen bonding interactions in DNA, methods of density functional theory at the B3PW91/6-311G** level have been used.

The presence of the three-centered H-bonds in DNA has been determined by studying the structural and energetic properties of a series of AA dimers. We have shown that the results are in good agreement with the experimental data. Four characterization indexes including 360° SAN rule, E(2) energy, Wiberg Bond Index (WBI) and spin-spin coupling constant have been applied and a set of characterization indexes has been successfully developed.

The results of this study revealed that the three-centered H-bonds are suggested to be present but not extending beyond (3.00Å) with minimum E(2) bond energy of 0.06 kcal/mol. Energetically, they are weaker than the regular intermolecular H-bonds, and hence, longer hydrogen bond distance is tolerated for the structural determination of three-centered H-bonds. Calculated results have shown that the E(2) energy of three-centered H-bonds has an exponential relationship with the bond distance. A small variation of bond distance in short distance region can lead to a significant reduction of bond energy. We have found that all seven three-centered H-bonds conform to the 360° SAN geometry rule. Hence, the four atoms involved in three-centered H-bonds formation are in the same plane in each dimer. The covalency of the three-centered H-bonds has been illustrated by the non-zero WBI values and the

J-coupling. The trend of the WBI values generally parallels with the trend of the E(2) energies. The non-zero spin-spin coupling constant ($^{1h}J_{XH}$) values further demonstrate that interaction between the hydrogen atom and the acceptor is present. The calculated negative $^{1h}J_{XH}$ values revealed that the three-center H-bonds have a poor electron-electron interaction due to the long bond distance with poor orbital overlapping.

We have also presented NBO analysis of the hydrogen bonds strength on the basis of interbase proton transfer in order to demonstrate if there is remote control effect on the whole cluster system. The results showed that the site of proton transfer can exert a certain effect of hydrogen bonds strength in both type of intermolecular and three centered H-bonds. We have demonstrated that the proton transfer system can be viewed as switches with remote communication ability on distance hydrogen bonds strength

.

Cooperativity effects is evident by increased E(2) energies of (AA)n dimer unit upon increasing the size of the oligomers. The sum of E(2) of hydrogen bonds in a dimer unit increased up to 10 % in maximum with n=5.

The results described herein are the first extensive set of calculation used to characterize the existence of three-centered H-bonds in DNA base pairs. These indexes could be use to predict the existence of three-centered H-bonds in a theoretical way which cannot be easily probed experimentally.

Armed with the knowledge acquired in this study, the role of solvent effect in three-centered H-bonds strength associated with proton transfer needs to be further

investigated. We believe, exploration of this aspect can yield new insights toward the understanding of DNA mutation mechanism.

REFERENCES AND APPENDIX

CHAPTER ONE

1.1 Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48-76.

1.2 Bhattacharyya, D.; Kundu, S.; Thakur, AR.; Majumdar, R. *J. Biomol Struct Dynam*, **1999**, *17*, 289-300. Rhodes, D. *Nucl Acids Res*. **1979**, *6*, 1805-1816.

1.3 Hobza, P.; Zahradnik, R. *Chem. Rev*. **1988**, 88, 871.
Gu,J.; Wang, J.; Leszczynski, J. *J. Am. Chem. Soc*. **2004**, *126*, 12651-12660

1.4 Steinter, T., *J. Chem. Soc. Perkin Trans*. **1995**, *2*, 1315.

1.5 Steiner, T.; Starikov, E. B.; Amado, A. M.; Teixeira-Dias, J. J. C., *J. Chem. Soc. Perkin Trans*. **1995**, *2*, 1321.

1.6 Jeffrey, G. A. *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, **1997**.

1.7 Emsley, J. *Chem Soc Rev* **1980,** *9*, 91-124.

1.8 Pimentel, G. C. and McClellan, A. L. *The hydrogen bond*, W. H. Freeman, San Francisco, **1960**.

1.9 Allen, F. H. *Acta Crystallogr. Sect. B* **1986**, *42*, 515-522.

1.10 Yanagi, K.; Prive, G. G.; Dickerson, R. E. *J. Mol. Biol.* **1991**, *217*, 201-214.

1.11 Parra, R. D.; Zeng, H.; Zhu, J.; Zheng, C.; Zeng, X. C.; Gong, B. *Chem. Eur. J.* **2001**, *7*,4352-4357.

1.12 Bureiko, S. F.; Glubev, N.S.; Pihlaja, K. *J. Mol. Struct.* **1999**, 297

1.13 Parra, R. D.; Zhu, J.; Zeng, X. C.; Gong, B. *J. Chem. Phys*. **2001**, *115*(13), 6030-6035.

1.14 Jeffrey, G. A. Saemger, W. *Hydrogen Bonding in Biological Structures*, Springer-Verlag, **1991**, pp. 139.

1.15 Taylor, R. *J. Mol. Struct*. **1981**, *71*, 311-325.

1.16 Espinosa, E.; Souhassou, M.; Lachekar, H.; Lecomte, C. *Acta Cryst.*, **1999**, *B55*, 563-572.

1.17 Louit, G.; P Hocquet, A.; Ghomi, M.; Meyer, M.; Suhnel, *J. Phys Chem Comm*, **2002**, *5*(15), 94-98.

1.18 Frank, H. S.; Wen, W. Y. *Discuss. Faraday Soc*. **1957**, *24*,133.

1.19 King, B. F.; Weinhold, F. *J. Chem. Phys*. **1995**, *103*, 333.

1.20 Kleeberg, H. *Interactions of water in ionic and nonionic hydrate; Kleeberg, H., Ed.; Springer-Verleg: Berlin-Heidelberg*, **1987**; p 89.

1.21 Kleeberg, H. *Interactions of water in ionic and nonionic hydrate; Kleeberg, H., Ed.; Springer-Verleg: Berlin-Heidelberg*, **1987**; p 113.

1.22 Kleeberg, H.; Klein, D.; Luck, W. A. *The Journal of physical Chemistry*, **1987**, *91*(12), 3200-3203.

1.23 Dannenberg, J. J.; Haskamp, L.; Masunov, A. *J. phys. Chem. A* **1999**, *103*, 7083-7086.

1.24 Suhai, S. *J. Chem. Phys.* **1994**, *101,* 9766.

1.25 Turi, L.; Dannenberg, J. J. *J. Am. Chem. Soc.* **1994**, *116*, 8714.

1.26 Bondi, A. *J. Phys Chem*, **1964**, *68*, 441-451.

1.27 Jeffrey, G. A. Saemger, W. *Hydrogen Bonding in Biological Structures*, Springer-Verlag, **1991**, pp. 23.

1.28 Zimmerman, S.C.; Murray, T. J. *Tetrahedron Lett*, **1994**, *35, 4077.*

1.29 Rozas, I. Alkorta, I. Elguero, J. *J. Phys. Chem. A.* **1998**, *102*, 9925.

1.30 Parthasarathy, R. *Acta crystallogr.* **1969**, *B25*, 509.

1.31 Bader, R. F. W. *Atoms in molecules. A Quantum Theory*, Clarendon, Oxford, **1990**.

1.32 Krivoruchka, I.G.; Vokin, A. I.; Turchaninov, V. K. *Russian Journal of Organic Chenistry*, **2002**, *38*(10), 1509-1514.

CHAPTER TWO

2.1 Sponer, J.; Leszczynski, J.; Hobza, P. *Journal of Bimolecular Structure & Dynamics*. **1996**, *14*,117-135.

2.2 Hehre, W. J. el al *A Brief Guide to Molecular Mechanics and Quantum Chemical Calculations*, Wavefunction, **1998**.

2.3 Burda, J. V.; Sponer, J.; Leszczynski, J. *Physical Chemistry Chemical Physics*, **2001**, *3*(19), 4404-4411.

2.4 Barfield, M.; Dingley, A. J.; Feigon, J.; Grzesiek, S. *J. Am. Chem. Soc.* **2001**, *123*, 4014-4022.

2.5 Vergenz, R. A.; Yazji, I.; Whittington, C.; Daw, J.; Tran, K. T. *J. Am. Chem. Soc.* **2003**, *125*, 12318-12327.

2.6 Kawahara, S.; Wada, T.; Kawauchi, S.; Uchimaru, T.; Sekine, M. *J. Phy. Chem. A*, **1999**, *103*, 8516-8523

2.7 Reed, A.E.; Curtiss, L.A.; Weinhold, F.; *Chem. Rev.* **1988**, *88*, 899

2.8 Weinhold, F. *Journal of Molecular Struture (Theochem)*, **1997**, *398-399*, 181-197

2.9 Pople, J. A.; McIver, J. W.; Ostlund, N. S. *The Journal of Chemical Physics*, **1968**, *49*, 2965-2970

2.10 Del Bene, J. E.; Perera, A. A.; Barlett, R. J. *J. Phys. Chem. A*, **2001**, *105*, 903-934

2.11 Borisova, N. P.; Semenov, S.G.; Stewart, J. J. P. *J. Chem. Soc. Dalton Trans.* **1973**, *838*, 2273

2.12 *Gaussian 98*, Revision D.3, Frisch, M. J.; Truck, G. W; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M, A,; Paghavachari, K.; A;-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B.

B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, **1998**.

CHAPTER THREE

3.1 Pauling, L. *The Nature of the Chemical Bond*, Cornell University Press, Ithica, **1942**, pp.192 From T7

3.2 Jeffrey, G. A.; Saemger, W. *Hydrogen Bonding in Biological Structures*, Springer-Verlag, **1991**

3.3 Nelson, H.C.; Finch, J.T.; Luisi, B.F. Klug, A. *Nature,* **1987**, *330*, 221-226

3.4 Parthasarathy, R. *Acta Crystallogr.* **1969**, *B25*, 509

3.5 Kwiatkowski, J. S.; Leszcynski, J. *J. Phys. Chem.* **1996**, *100*, 941

3.6 Estrin, D.A.; Paglieri, L.; Corongiu, G.. *J. Phys. Chem.* **1994**, ***98***, 5653

3.7 Sponer, J.; Leszczynski, J.; Hobza, P.; *J. Biomol. Struct. Dyn.*, **1996**, ***14***, 117

3.8 Weinhold, F. *Journal of Molecular Structure*, **1997**, *398-399*, 181-197

3.9 Dingley, A. J.; Cordier, F.; Grzesiek, S. *Concepts in Magnetic Resonance*, **2000,** *13*, 103-127,

3.10 Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *Journal of the American Chemistry Soceity*, **2000**, *122*, 3560-3561

3.11 Dingley, A.J., Masse, J.E., Feigon, J., and Grzesiek, S. *J. Biomol. NMR*, **2000**, *16*, 279–289

3.12 Wang, Z. X.; Schleyer, P. V. R. *Angrew. Cehm. Int, Ed.* **2002**, *41*, 4082

3.13 Reed, A. E., Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899

3.14 Ly, A.; Tran, N. Q.; Ward, J. F.; Milligan, J. R. *Biochemistry*, **2004**; *43*(28); 9098-9104.

3.15 King, B.F.; Weinhold, F. *J. Chem. Phys*. **1995**, *103*, 333-347

3.16 Peralta, J. E.; Ruiz de Azua, M. C.; Contreras, R. H. *J. Mol. Strct*, **1999**, *491*, 23-31

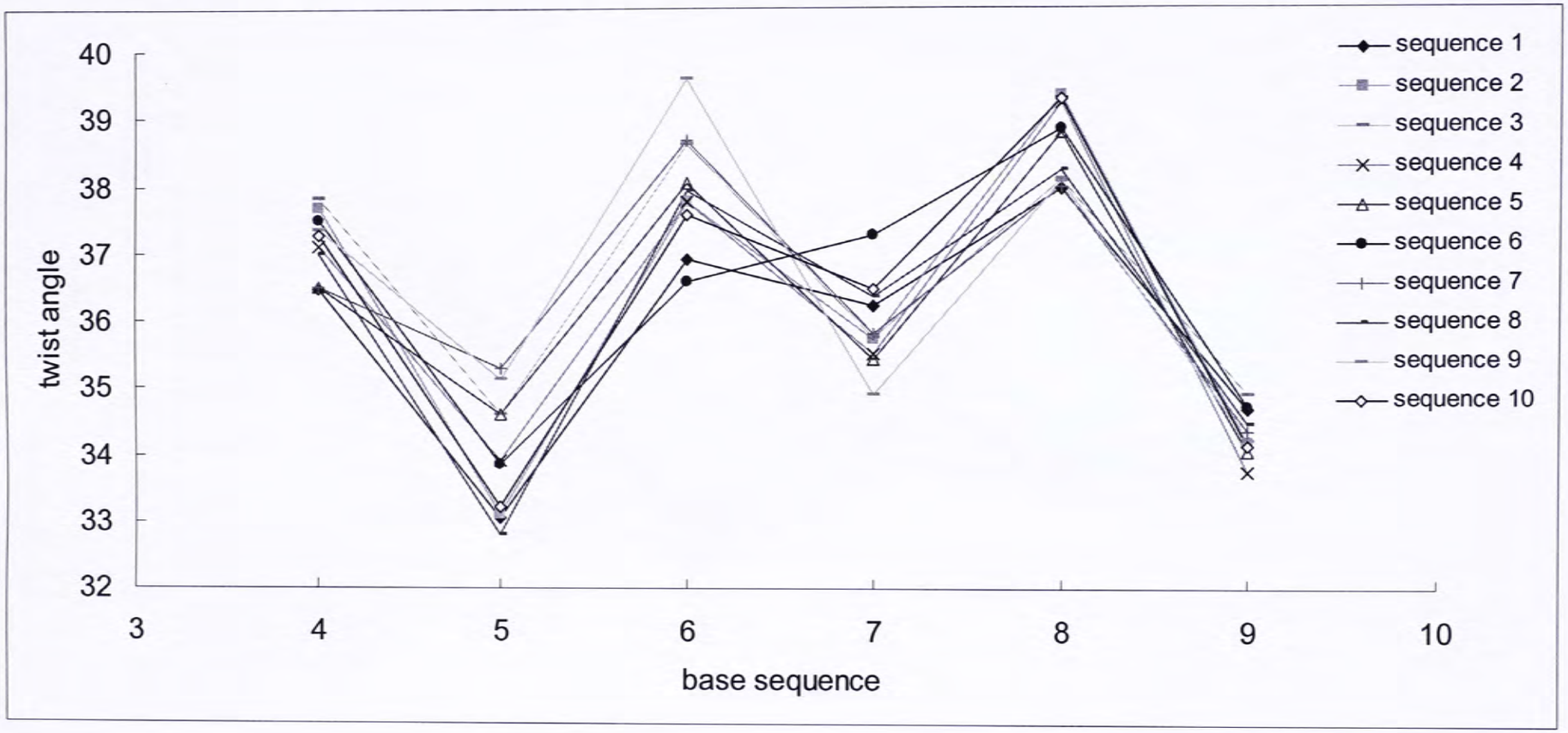
Supplementary Table 1 Deletion energy of the intermolecular and three-centered hydrogen bonds between H and X across two base pair

A8A9 dimer	Three-centered H-bond		Deletion energy (kcal/mol)		Distance (Å)	
	Donor	Acceptor	E(H…B)	E(H…B')	r(H…B)	r(H…B')
A	A8(N6)-A8(H6a)	T16(O4)	3.328	1.314	2.25	2.33
B	A8(N6)-A8(H6a)	T16(O4)	1.463	0.883	2.17	2.45
C	A8(N6)-A8(H6a)	T16(O4)	4.053	0.650	2.21	2.53
D	A8(N6)-A8(H6a)	T16(O4)	3.578	0.630	2.23	2.58
E	A8(N6)-A8(H6a)	T16(O4)	5.551	0.439	2.12	2.64
F	A8(N6)-A8(H6a)	T16(O4)	17.886	0.126	1.79	2.81
G	A8(N6)-A8(H6a)	T16(O4)	-	0.058	1.65	2.95

Supplementary Table 2 Intermolecular and three-centered hydrogen bond energies of A8A9, A9C10 dimers in different size of clusters

Hydrogen bonds	E(2) energy			r(H…B)
A-H…B	dimer (A8A9)	dimer (A9C10)	trimer	
A8-T17 base pair				
N6-H6a···O4	4.24		4.24	2.17
N1···H3-N3	33.26		33.14	1.69
C2-H2···O2	0.82		0.82	2.44
A9-T16 base pair				
N6-H6a···O4	1.36	1.68	1.66	2.38
N1···H3-N3	31.24	30.12	30.94	1.69
C2-H2···O2	0.68	0.68	0.68	2.45
C10-G15 base pair				
N4-H4a···O6		3.94	3.92	2.02
N3···H1-N1		23.99	23.97	1.73
O2···H2a-N2		29.01	29.06	1.61
A-H…B'				
T16(O4)···A8(H6a)-A8(N6)	0.88		0.88	2.45
A9(N1)···G15(H1)-G15(N1)		0.04	0.04	3.16
G15(O6)···A9(H6a)-A9(N6)		0.92	0.92	2.12

Supplementary Table 3 Twist angle distribution of each base of d(5-G1 G2 C3 A4 A5 A6 A7 A8 A9 C10 G11 G12-3')$_2$ sequence



Supplementary Table 4 Rise angle distribution of each base of d(5-G1 G2 C3 A4 A5 A6 A7 A8 A9 C10 G11 G12-3')$_2$ sequence